

## Two distinct forms of the 64,000 $M_r$ protein of the cleavage stimulation factor are expressed in mouse male germ cells

A. MICHELLE WALLACE\*, BRINDA DASS\*, STUART E. RAVNIK\*, VIJAY TONK†, NANCY A. JENKINS‡,  
DEBRA J. GILBERT‡, NEAL G. COPELAND‡, AND CLINTON C. MACDONALD\*§¶

Departments of \*Cell Biology and Biochemistry and †Pediatrics and Pathology and §Southwest Cancer Center at University Medical Center, Texas Tech University Health Sciences Center, Lubbock, TX 79430; and ‡Mammalian Genetics Laboratory, Advanced BioScience Laboratories/Basic Research Program, National Cancer Institute/Frederick Cancer Research and Development Center, Frederick, MD 21702

Communicated by Thomas E. Shenk, Princeton University, Princeton, NJ, April 5, 1999 (received for review February 3, 1999)

**ABSTRACT** Polyadenylation in male germ cells differs from that in somatic cells. Many germ cell mRNAs do not contain the canonical AAUAAA in their 3' ends but are efficiently polyadenylated. To determine whether the 64,000  $M_r$  protein of the cleavage stimulation factor (CstF-64) is altered in male germ cells, we examined its expression in mouse testis. In addition to the 64,000  $M_r$  form, we found a related ~70,000  $M_r$  protein that is abundant in testis, at low levels in brain, and undetectable in all other tissues examined. Expression of the ~70,000  $M_r$  CstF-64 was limited to meiotic spermatocytes and postmeiotic spermatids in testis. In contrast, the 64,000  $M_r$  form was absent from spermatocytes, suggesting that the testis-specific CstF-64 might control expression of meiosis-specific genes. To determine why the 64,000  $M_r$  CstF-64 is not expressed in spermatocytes, we mapped its chromosomal location to the X chromosome in both mouse and human. CstF-64 may, therefore, be absent in spermatocytes because the X chromosome is inactivated during male meiosis. By extension, the testis-specific CstF-64 may be expressed from an autosomal homolog of the X chromosomal gene.

Polyadenylation, the process of 3' end formation in eukaryotic mRNAs, is required for synthesis, transport, translation, and stability of most mRNAs (1–4). The sequence AAUAAA specifies accurate and efficient addition of poly(A) to the mRNA 3' end (5, 6). Substitutions at any position in this sequence diminish polyadenylation *in vivo* and *in vitro* (7–9), and 90–95% of sequenced mRNAs have AAUAAA in their 3' ends (1, 2). However, many mRNAs expressed in male germ cells do not have AAUAAA (refs. 10–13 and Table 1), but are nevertheless efficiently polyadenylated. Sequences such as UAUAAA, AUUAAA, UACAAA, and GAUAAA might substitute for the normal polyadenylation signal (10, 11), but no experimental evidence support these assertions.

One hypothesis to account for the use of non-AAUAAA signals is that a protein involved in polyadenylation is altered in germ cells. A candidate is the 64,000  $M_r$  subunit of the cleavage stimulation factor (CstF-64) (14, 15), one of three polypeptides of CstF (16). CstF-64 has been shown to be essential for growth and viability of avian cells (17) and acts by binding to a U- or G/U-rich region downstream of the cleavage site (18) whereas a 160,000  $M_r$  protein of the cleavage and polyadenylation specificity factor binds to the AAUAAA signal (19, 20). CstF-64 governs polyadenylation site choice in adenovirus (21) and is involved in the immunoglobulin switch from a membrane to secretory form during B-cell maturation (22–24). CstF also is involved in the cooperation of polyadenylation with splicing (25), and recent results show that it interacts with the C-terminal domain of RNA polymerase II and thus couples polyadenylation with transcription (26).

Because of its essential role in somatic cell polyadenylation, we examined CstF-64 expression in male germ cells to test the hypothesis that it might contribute to polyadenylation of non-AAUAAA-containing mRNAs. We found that two different proteins immunologically related to CstF-64 are expressed in testis, and to a lesser extent in brain, whereas a single form is expressed in all other tissues. We found that, of the two forms, the somatic form is expressed in germ cells before and after, but not during meiosis. In contrast, the testis-specific form is expressed exclusively during meiosis and afterward. Mapping of the CstF-64 gene to the X chromosome in both mouse and human suggests that the somatic form of CstF-64 is absent in meiotic cells because of X chromosome inactivation. It further suggests that a second, autosomal gene encodes the testis-specific CstF-64. Exclusive, stage-specific expression of an essential mRNA processing protein during meiosis could be important in the global regulation of gene expression in those cells.

### MATERIALS AND METHODS

**Protein Analysis.** HeLa nuclear extract was made as described (29). Testicular tubules were prepared by decapsulating freshly harvested mouse testes in PBS and washing to remove interstitial cells. Nuclei were prepared by centrifugation over a sucrose pad as described (30), were boiled and sonicated in loading buffer (31), were separated by 10% SDS/PAGE, and were transferred to PVDF-Plus (Micron Separations, Westborough, MA) (32) for immunoblotting. Membranes were stained with Ponceau S and were incubated for 1–16 hours in block solution (TBS with 10% nonfat dry milk). Monoclonal antibodies 3A7 and 6A9, specific for the human CstF-64, are described elsewhere (16) and were used at a dilution of 1:50. Membranes were incubated with primary antibody in block followed by incubation with horseradish peroxidase-coupled sheep anti-mouse IgG (1:10,000). After washing in TBS, immunoreactive bands were visualized by chemiluminescence by using the Pierce SuperSignal kit.

**Immunohistochemistry.** Immunohistochemistry of paraformaldehyde-fixed paraffin-embedded sections of adult mouse testis was done according to Ravnik and Wolgemuth (12) by using the Vectastain ABC kit (Vector Laboratories). Sections were blocked with 1 × PBS/0.1% Triton X-100/2.5% normal horse serum for two hours and were incubated with primary antibody at 1:100 in 1 × PBS/0.1% Triton X-100/1.5% horse serum for 16–18 hours at 4°C. After washing, sections were incubated with secondary biotinylated antibody (anti-mouse IgG, 1:200) and were incubated with avidin-biotin complex

Abbreviations: CstF, cleavage stimulation factor; FISH, fluorescence *in situ* hybridization.

\*To whom reprint requests should be addressed at: Department of Cell Biology and Biochemistry, Texas Tech University Health Sciences Center, 3601 Fourth Street, Lubbock, TX 79430. e-mail: cbccm2@ttuhsc.edu.

according to the manufacturer's directions. Antibody staining was developed in 0.1 M Tris-HCl (pH 7.2) with 0.2 mg/ml of 3, 3' diaminobenzidine. Slides were counterstained by using Harris' Hematoxylin and were viewed with an Olympus (New Hyde Park, NY) BX-60 photomicroscope.

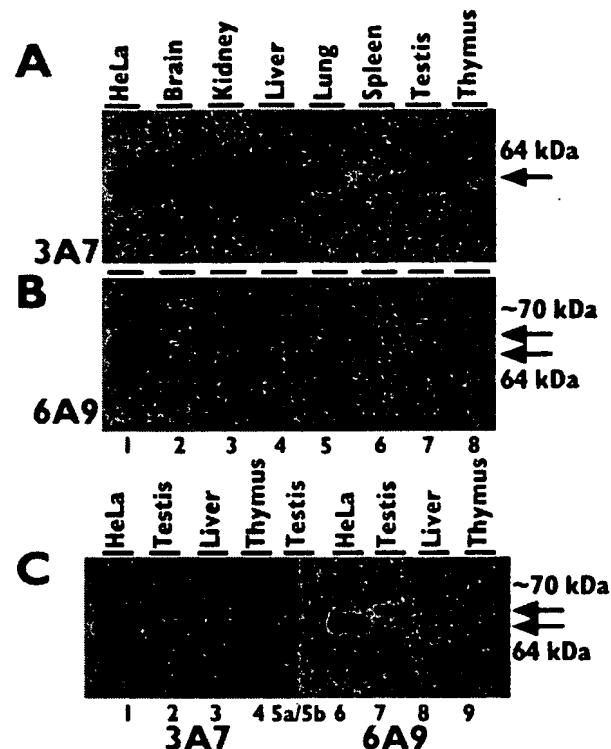
**Fluorescence in Situ Hybridization (FISH).** A P1 artificial chromosome containing the human genomic CstF-64 [dJ347M6 (33), GenBank accession no. Z95327], provided by Gareth Howell (The Sanger Centre, Cambridgeshire, U.K.), was labeled with digoxigenin-11-dUTP (Boehringer Mannheim) by using the Oncor large fragment labeling kit (Oncor). Probe was hybridized with human Cot-1 DNA (GIBCO/BRL), was resuspended in hybridization solution [50% formamide/10% dextran sulfate/2× standard saline citrate (SSC)], and was denatured at 80°C before use. Metaphase spreads from human peripheral white blood cells (34) were denatured in 2× SSC and 70% formamide at 72°C, were dehydrated in an ethanol series, and were air dried. Chromosomes were hybridized at 37°C for 12–16 hours. Washes were for 5 min at 65°C in 2× SSC (pH 7.0) followed by 1× phosphate-buffered detergent (Oncor). Slides then were incubated with FITC-antidigoxigenin antibody for 45 min at 37°C, were washed in 1× phosphate-buffered detergent, and were mounted in Antifade (Oncor) with propidium iodide counterstain. Chromosomes were visualized by using a Zeiss Axioskop microscope with epifluorescence using FITC and propidium iodide single-band pass filters. Images were captured by using a charge-coupled device camera and a MACKTYPE 5.4 imaging system (Perceptive Scientific Instruments, League City, TX) with the 100×, 1.3 numerical aperture objective under oil.

**Interspecific Mouse Backcross Mapping.** Interspecific backcross progeny were generated by mating (C57BL/6J × *Mus spreitus*) F1 females and C57BL/6J males as described (35). A total of 205 N<sub>2</sub> mice were used to map the *Cstf64* locus. Southern blot analysis was performed as described (36) with Hybond-N+ nylon membranes (Amersham Pharmacia). The probe, a 342-bp fragment of mouse CstF-64 cDNA from the 3' untranslated region [nucleotides 1809–2151 (B.D. and C.C.M., unpublished work)], was labeled with [ $\alpha$ -<sup>32</sup>P]dCTP by using a random primed labeling kit (Stratagene); washing was done to a stringency of 0.5× SSCP (120 mM NaCl/5 mM sodium citrate/20 mM sodium phosphate, pH 6.8) and 0.1% SDS at 65°C. A fragment of 7.9 kilobases was detected in *TaqI*-digested C57BL/6J DNA, and a fragment of 6.0 kilobases was detected in *M. spreitus* DNA. Presence or absence of the 6.0-kilobase *TaqI* *M. spreitus*-specific fragment was followed in backcross mice. Probes and restriction fragment length polymorphisms for the loci linked to *Cstf64* have been reported (37). Recombination distances were calculated by using MAP MANAGER 2.6.5 (Roswell Park Cancer Institute, Buffalo, NY; <http://mebio.med.buffalo.edu/mapmgr.html>). Gene order was determined by minimizing the number of recombination events required to explain the allele distribution patterns.

## RESULTS

**Two Forms of CstF-64 Are Expressed in Mouse Testis.** Monoclonal antibodies to human CstF-64 were used to examine CstF-64 in mouse tissues (Fig. 1). Protein immunoblots of nuclear extracts from various mouse tissues were incubated with the monoclonal antibodies 3A7 and 6A9 (16). The 3A7 antibody detected a protein of  $\sim$ 64,000  $M_r$  in HeLa cell (Fig. 1A, lane 1) and all mouse tissues examined (lanes 2–8), corresponding to CstF-64 first described in HeLa cells (14, 16).

With the 6A9 antibody, a 64,000  $M_r$  reactive protein was detected in HeLa nuclear extract (Fig. 1B, lane 1). In contrast, a protein of  $\sim$ 70,000  $M_r$  was detected in mouse testicular extracts (Fig. 1B, lane 7) and in small amounts in brain (lane 2). No protein was detected in other tissues (Fig. 1B, lanes 3–6 and 8). To ensure that the  $\sim$ 70,000  $M_r$  protein detected with



**FIG. 1.** Protein immunoblot detection of CstF-64 in mouse tissues. (A and B) HeLa nuclear extract (lane 1), nuclei from indicated mouse tissues (lanes 2–6 and 8), or whole testicular tubules (lane 7) were examined by protein immunoblotting. Testis tubules were used rather than nuclei because of the difficulty in obtaining germ cell nuclei (66). Immunoblots were incubated with 3A7 (A) or 6A9 (B) monoclonal antibodies. (C) A single immunoblot membrane was split along one lane (5a/5b), was incubated with 3A7 (Left) or 6A9 (Right), and was reassembled. Lanes contain HeLa nuclear extract (lanes 1 and 6), whole testicular tubules (lanes 2, 5a/5b, and 7), liver (lanes 3 and 8), or thymus (lanes 4 and 9) nuclei. Arrows indicate approximate sizes of CstF-64 proteins.

6A9 did not comigrate with that detected by 3A7, we repeated immunoblots with 3A7 and 6A9 but bisected a lane containing testis extracts (Fig. 1C, lane 5a/5b). After realignment of the filters, it was evident that the protein detected with 6A9 had a substantially different mobility than that detected with 3A7.

**The  $\sim$ 70,000  $M_r$  CstF-64 Is the Only Form Expressed in Spermatocytes.** Expression of 3A7-reactive (64,000  $M_r$ ) and 6A9-reactive ( $\sim$ 70,000  $M_r$ ) CstF-64 were examined immunohistochemically in mouse testis sections. The 3A7 antibody detected immunoreactivity in nuclei of testicular somatic cells, primarily Leydig, Sertoli, and macrophage cells (Fig. 2A) and detected protein in the nuclei of kidney, liver, thymus, and epididymal cells (data not shown). This is consistent with 3A7 immunoblot data showing the 64,000  $M_r$  form in every tissue examined.

Within the seminiferous epithelium, different germ cell types were characterized by their association with each other and with types found in different cross sections and were categorized into 12 stages of the seminiferous epithelium (38). 3A7 detected strong immunoreactivity in all spermatogonia and early spermatocytes. The intensity of 3A7 staining diminished in pachytene spermatocytes in stage II–IV tubules and was not detected again until after meiosis in round spermatids of stage IV tubules (Fig. 2A and summarized in Fig. 2E). Spermatid staining with 3A7 remained until spermatid elongation (stage X) at approximately the same time as when protein translation in spermatids ends (see Fig. 5).

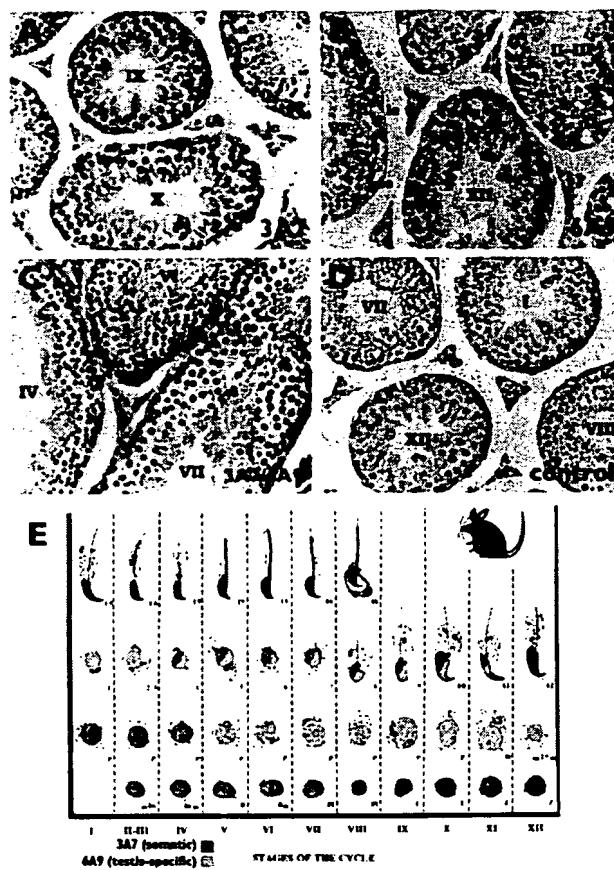


FIG. 2. Immunohistochemical localization of CstF-64 in mouse testis sections. Sections of mouse testis were incubated with 3A7 (A), 6A9 (B), a mixture of 3A7 and 6A9 (C), or an irrelevant control antibody, 2A6 (D). Roman numerals refer to the stage of the seminiferous epithelium. lc, Leydig cell; sc, Sertoli cell; sg, spermatogonia; spc, spermatocyte; rs, round spermatid; es, elongating spermatid. (×100.) (E) Summary of CstF-64 protein expression in mouse germ cells (see also Fig. 5). 3A7 (somatic form) reactivity is in red; 6A9 (testis-specific) is in blue. The diagram is adapted from Russell *et al.* (38).

In marked contrast, 6A9 recognized an epitope beginning in stage V pachytene spermatocytes and continuing in all spermatocytes to early spermatids (Fig. 2B, summarized in Fig. 2E). Expression continued in spermatocytes through meiosis and terminated in elongating spermatids at stage X. No staining with 6A9 was detected in early germ cells or any somatic cells (Fig. 2B), including those from the liver, thymus, and kidney (data not shown). When testis sections were incubated with a mixture of both 3A7 and 6A9, staining was observed in all testicular cell types except late-elongating spermatids and spermatozoa (Fig. 2C), suggesting that no cell type in the testis lacks one or the other form of CstF-64. The 2A6/E1B antibody is of the same isotype (IgG<sub>1</sub>) as 3A7 and 6A9 and was used as a control (16). This antibody displayed no appreciable nuclear or cytoplasmic staining of any testicular cell type (Fig. 2D).

**Mapping of CSTF64 to the Human X Chromosome.** The P1 artificial chromosome dj347M6 contains the CstF-64 genomic clone and maps to the long arm of the human X chromosome (Xq22). This genomic clone is entirely colinear with the human cDNA, as confirmed by sequence comparison and Southern hybridization (data not shown).

Further confirmation of X chromosome assignment of CSTF64 was obtained by FISH analysis of human metaphase spreads by using dj347M6 (see *Methods and Materials*). Over 150 spreads were examined in samples from females, and 75

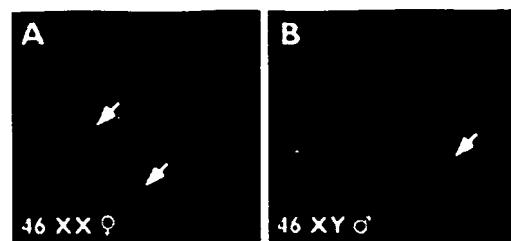


FIG. 3. FISH analysis of human metaphase chromosomes. Metaphase spreads from peripheral blood were prepared as described in the text. Hybridization was with dj347M6 containing the human CstF-64 gene. (A) Representative metaphase spread from a female. (B) Metaphase spread from a male. Yellow arrows indicate X chromosomal signals from probe hybridization. The blue arrow indicates signal in an intact interphase cell.

were used from males. FISH showed two high-intensity signals from females (Fig. 3A) and one from males (Fig. 3B) on medium-sized (Group C) chromosomes that are either metacentric or submetacentric, consistent with X-chromosomal localization of CSTF64. Several lower-intensity signals also can be seen at other locations in interphase cells (Fig. 3B).

**Cstf64 Is X Chromosomal in Mouse, and Cstf64-Related Genes Are on Other Chromosomes.** The mouse chromosomal location of *Cstf64* was determined by interspecific backcross analysis using progeny derived from matings of [(C57BL/6J × *M. spretus*) F1 × C57BL/6J] mice. This panel has been typed for >2,800 loci that are distributed among all autosomes as well as the X chromosome (35). C57BL/6J and *M. spretus* DNAs were digested with several enzymes and were analyzed for informative restriction fragment length polymorphisms by using a mCstF-64 3' untranslated region probe. The 6.0-kilobase *TaqI* *M. spretus* restriction fragment length polymorphism (see *Materials and Methods*) was used to follow segregation of the *Cstf64* locus. Results indicated that *Cstf64* is located in the central region of the X chromosome linked to *Pgk1* and *Btk*. One-hundred and thirty-five mice were analyzed for every marker (Fig. 4), and up to one-hundred and eighty-five mice were typed for some pairs of markers. Each locus was analyzed for recombination frequencies by using the additional

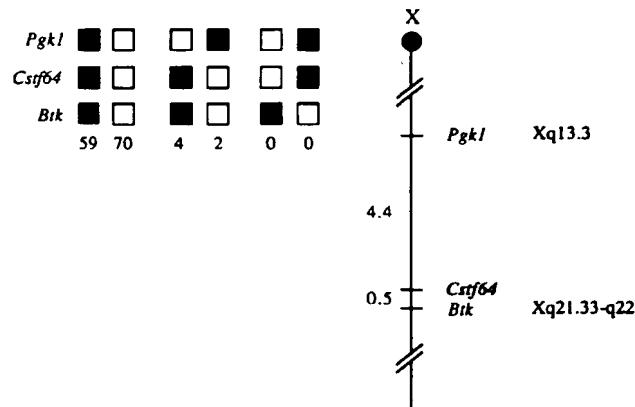


FIG. 4. Cstf64 maps in the central region of the mouse X chromosome. (Right) Columns represent the chromosome identified in backcross progeny that inherited from the (C57BL/6J × *M. spretus*) F1 parent. Shaded boxes represent a C57BL/6J, and white boxes represent a *M. spretus* allele. The number of offspring inheriting each type of chromosome is listed at the bottom. (Left) Partial X chromosome linkage map showing *Cstf64* in relation to linked genes. Recombination distances between loci are shown to the left, and positions of loci in human chromosomes are shown to the right. Human loci cited can be obtained from Genome Data Base, a computerized database maintained by The William H Welch Medical Library of The Johns Hopkins University (Baltimore).

Table 1. Non-AAUAAA polyadenylation signals in germ cell transcripts

Ref.	Ref.	Sequence
Blast results		
Bovine sex-determining region Y (SR Y) (site 1)		TTCTTGTGCTTATTTCATAATTGACTTCCITTACTCTCGCTAACAAAGGGCCTCTTATCTCAATTTTAC
Bovine sex-determining region Y (SR Y) (site 2)		ACCTGCGACTTAATTTTATAAGAAGCTAACATGTAAGAATTAGACTTTCTGCTAACACTTTCC
Bovine sex-determining region Y (SR Y) (site 3)		GCCAAAGAAGTAAAGTACATTAAACATGTAAGAATTAGACTTTCTCAATAAACTGGCCCTCTGTTAATC
Bovine sex-determining region Y (SR Y) (site 4)		TAACCTGGTCCTCTGTGTTAATCAGTTCTTCTCAATAAGAGTACTTTTGTAAATTATCTTAACAGCAC
Human spermine synthase		ATTAAATGGTGTGCTTAAAGAGAGGTTAAAAAATAGGCTGTTCAATTAAATTACAAAGAG
Human cyclic AMP-dependent protein kinase RII $\alpha$		ACGGCATGTCCTCCTCAACAAATTAAAGTGCCTGATACAAAGTCCAAAGTATAAACATGCTCCTTCTCTC
Rat cyclic AMP-dependent protein kinase RII $\beta$		TTCAATATGGTTATAATATGAGCATACATGTTTAAGAAGATAATTAAAGATGTACTCATAGGCC
Mouse androgen receptor		ATAAAACAATCTAAATTTCAGATGAAATGTCATCAGAGTTCTTTCATAAGCTTATTAAATTTGAAGATACTT
Mouse calactin		CTTCTTCCTGCCTTCTCTTGTCTTGGTTGGCTTCTAAGCTCTCTAGATCAATTAAATTTGGCTGTCT
Mouse germ cell-specific lamin B $_3$		CTGTATGTTGAAAGAACCTGCCCCCAGGGTCACCCATGGTCCCTAGGCCCTAAAAACTCTGAAACCTT
Mouse mRNA calmein (meg 1)		AAAACAAATTCAAAATTCTGCTTTAATTCAATTGTTCTTACACAAACATAAAACCCACCTGAAACCTT
Mouse Sp17 gene for sperm specific protein		TGAGAATCTGAAAGAAGGGCAGACAATTGAGACCACAGGTTAACCCCGAAACATGAAAGTAATC
Mouse testis-specific c-abl oncogene		ACTGGTTTTCATGGAAAGGAAGCTGCTTACTGTACCTGACCTTGTACCTGTCACAAACTGTC
Mouse transition protein 2		GATGTCACAGGAAACATGTCACTGTCAGCAAAATTCTATGCAACATGGATAAAGCTGTTGACCTG
Mouse tyrosine phosphatase PTP-1		AGAGGAAATGCCAAATACTCTCTGCTCTCACCTCACTCCTCTTCTGATTCAATTGTTTTGG
Mouse CREM $\tau$ (site 1)		CTATGAACTGATTACAGCATGTACAGTGGTCAATGCAATAATATAGCCGGCAAGAAATTAG
Mouse CREM $\tau$ (site 2)		TATTGCTGTCCTTGGGTGCCCCAACAGTGTGAAAGACAGTTTAATAGAACCCATGTTGCTCTTGTACTA
Mouse CREM $\tau$ (site 3)		GGGATTACGGATGTGGGCCAGAAATATCCAGTTGATCAAGTTCTTATTAATTAATTACTTCTTTTA
Mouse cyclin A2 (site 2)		GAAACTAACTTGATTGTTTATTGTTACAGTGTGGGAGTAGGAAAGGATTTTAAATAAATATTCA
Mouse cyclin A2 (site 3)		AGCTCTTGCACTCTGTGAAAATTGTGATAAGCTTCTATAAAACTCTTCTTCTTCTTCTTCTTCTT
Mouse GCNFRTR		ACACATTTGTCGTGTTAGAGAGAACACTGCCTTTTATAGCCACAGCTTAAGCTTGTGTCACGGT
Mouse histone H2a		CCCCCGAGCTCCAAAAGGCTCTTTCAGGCCCCACTGATCAGTAAAGGTTATGACGTTT
Mouse histone H3		CAGAGCCACCCCAATTTCTATAAAGGCTGTATATGCAATAAGCTTCTTATAAAACCCCACTCAGCAACTCC
Mouse mRNA for zinc finger protein		CGCCCTGAGGAAAATGTTCTGTTGTTGGTAGGTCTTTTAAATTTTTTTTAAATTTTTTTTTT
Mouse seip POU-domain protein		AATTAAATCGCTGTATGCTGTTATTATCGGGGCCCTCCCCCCTCCAGAGGGGGGGCGAGAGCCGCATCTGTGTT
Mouse Zfp-29 zinc finger protein		CAACCCCTCCCCCAACCCCCCTTGGTCTCCCTTAAAGGAAACCCTTCTCCTTAAT
Mouse Zfp-35 zinc finger protein		GTACTGTGTTAGTATAAAAAGGCTATGAGAAATAACACTCACTGTCGGGGTAAATGACCTGCAAGCCCTCCC
Mouse Zfp-59 zinc finger protein		AAAAAAAGAAGGGAATGAACAGCATGGATAATAAAGGTGTCATTGTCAGGGTCAAGGTCAACTCATGG
Mouse phosphoprotein phosphatase mRNA		GTAAATCAGTCGCAATTATTCTGTAAGAAAATGACCATTTATAAACTTCTTAATTATGTTCAATT
Mouse B94, inducible by tumor necrosis factor		GGGACATCAACACGGGGAGTGCAGGGCCTCCAGGCCCTTGTCTCCCTATBAGGTTACTTAGGTT
Mouse sp56, sperm fertilization protein		GCCATGTTGTTCAATTGTCATTGTTACTTGGCTGGACTCATGTAATAGTTAAACTCTGTCAAATGTCA
Pig zonadhesin		GAGGAGGGCTAACAAAGTGGCTTGTGAGGTGTTCAACACAAAGGGAGAGATAAAATTATTTTATT
Mouse zonadhesin		AGAGCTTACCTTCAAGGCCACCGAGTTCTGAGCTGTCTCCTCAAAAGAGAAAAATTATTATTATT
Calcineurin B $\beta$ 1		ACCCAACACTTTTGCCTTCTCCATCTCTGTAAGATCTGCTCAAGACGTCAGCTCTGTGTT
Calcineurin B $\beta$ 2		CTTCCCTGTGGCATTCAAGACTTTAACAGATTGTTAACCTCTATGTTGAAACTAGC
Rat <i>pim-1</i> protooncogene		CTGTGGTCACCAAGATGGCAGGGTAGGGAAAAACATTTGGGGGGGATTAACAAACTAGCACC
Rat calmodulin-binding protein		ATGATAAGCACTAATGTTGATTCTGTTGAGGTGTTAACATATACAGTGGTAAACCT
Rat Tegf (tc2)		CTTCTCCCTCAGGCCCGTTCTTGCACACATCACAGGTGTCGGTCCATGAATGAAAGTGAATG
Rat TPI-2		GGGGCGACTGAGGGGGCAAGGGTAAAGGGTAAAGGATTAAGGATCAACAGGGTAAAGGGTAAAGG

The 3' ends of germ cell-expressed mRNAs were surveyed. The last nucleotide on the right indicates the site of poly(A) addition as cited. Underlined hexanucleotides indicate putative polyadenylation signals for the mRNA. Shown are 39 cDNAs for which the site of poly(A) addition could be determined but that lacked a discernible AAUAAA sequence within 70 nucleotides of the 3' end.

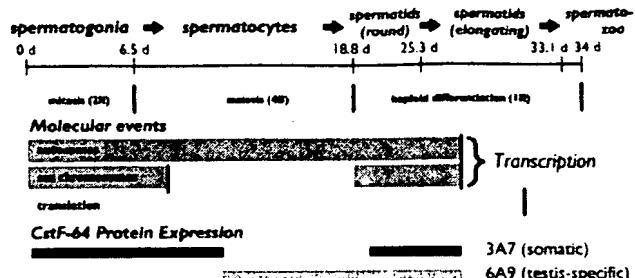


FIG. 5. Representation of essential events in mouse spermatogenesis with emphasis on sex chromosome transcription and CstF-64 expression. Stage and times (in days) of spermatogenesis are shown at top. Approximate durations of molecular events (Middle) and CstF-64 protein expression (Bottom) are expressed as bars.

data. Ratios of the number of mice with recombinant chromosomes to the number analyzed for each pair of loci and the most likely gene order are: centromere—*Pgk1*—6/135—*Cstf64*—1/185—*Btk*. The recombination frequencies (genetic distances in centimorgans  $\pm$  the standard error) are *Pgk1*—4.4  $\pm$  1.8—*Cstf64*—0.5  $\pm$  0.5—*Btk*. Hybridization with a full-length cDNA probe detected multiple restriction fragments, at least one of which is autosomal (data not shown). These results suggest that there are several *Cstf64*-related genes in the mouse genome.

## DISCUSSION

We have shown that two distinct forms of the essential polyadenylation protein CstF-64 are expressed in mouse testis. The 64,000  $M_r$  somatic form, detected with the 3A7 monoclonal antibody, was found in all tissues examined and is undoubtedly orthologous to CstF-64 described in HeLa cells (16). An  $\sim$ 70,000  $M_r$  form, detected with the 6A9 antibody, was seen only in extracts from testis and brain. We also have shown that the CstF-64 gene is encoded by a locus on the X chromosome in both mouse and human. In mouse, *Cstf64*-related genes are present on other chromosomes.

We initially were surprised that we saw two forms of CstF-64 in mouse tissues using antibodies that recognize only a single form in HeLa cells. However, subsequent experiments have shown that the human cDNA for the somatic CstF-64 encodes determinants for both the 3A7 and 6A9 epitopes (B.D., K. W. McMahon, and C.C.M., unpublished work). In these same experiments, the mouse somatic cDNA encodes only the 3A7 epitope. Presumably, a similar analysis of the cDNA for the mouse testis-specific CstF-64 would have only the 6A9 epitope. This fortuitous segregation of the two epitopes has made it possible for us to examine the two forms individually in mouse, which would not be possible in human.

Immunohistochemical results (summarized in Figs. 2E and 5) established expression of the somatic form of CstF-64 (3A7) in spermatogonia and early spermatocytes. In contrast, the testis-specific CstF-64 (6A9) was the only form expressed during most of meiosis. Subsequently, in round spermatids and early elongating spermatids, both the somatic and testis-specific forms were expressed. This suggests the possibility of three phases for polyadenylation in germ cells: phase I (pre-meiosis), in which normal polyadenylation is mediated by the somatic form of CstF-64. Phase II (meiosis) would be mediated by the testis-specific form. Phase III (postmeiosis) would be mediated by the simultaneous presence of both the somatic and testis-specific forms of CstF-64. This also suggests the possibility that the testis-specific form of CstF-64 is responsible for use of non-AAUAAA polyadenylation signals in germ cell mRNAs during and after meiosis (see Table 1).

The functional significance of these unconventional polyadenylation signals in male germ cells is not clear. Many mRNAs are expressed in multiple tissues, but the testicular

message uses a different polyadenylation site (see Table 1). For example, the mRNA for the *c-abl* oncogene uses a different site during germ cell maturation (10), as does *pim-1* (39), CREM $\tau$  (40, 41), and cyclin A2 (12). In each case, the polyadenylation mechanism overlooks a canonical AAUAAA polyadenylation signal for a germ cell-specific site, yet these mRNAs have normal poly(A) tails.

FISH analysis in human and interspecific backcross analysis in mouse maps the *Cstf64* gene to the X chromosome in both species (Figs. 3 and 4). We compared our map of the X chromosome with a mouse linkage map of uncloned mouse mutations (Mouse Genome Database, The Jackson Laboratory). *Cstf64* is in a region that lacks mutations with a phenotype that might be expected for an alteration in this locus. However, this region shares homology with the long arm of the human X chromosome (Fig. 4), consistent with the finding that the human homolog of *Cstf64* maps to Xq.

CstF-64 is a phosphoprotein (refs. 22 and 42; C.C.M., unpublished work). However, the difference in mobility of the  $\sim$ 70,000  $M_r$  and the 64,000  $M_r$  forms likely are not a result of phosphorylation; our experiments with one- and two-dimensional electrophoresis show that the  $\sim$ 70,000  $M_r$  form does not comigrate with any phosphorylated form of somatic CstF-64 (data not shown). Therefore, three mechanisms might account for different protein forms of CstF-64 in germ cells: (i) the testis-specific protein could be the product of a different gene; (ii) it could be a post-translational modification of the primary protein that does not involve phosphorylation; or (iii) it could be an alternatively spliced product of the mRNA. Although we have not completely eliminated the second or third possibilities, we believe our data strongly support the first. Specifically, because the gene for the somatic CstF-64 is on the X chromosome, transcription of that gene is likely inactivated during male meiosis (28). For inactivation of essential X-linked genes, often an intronless retroposon on an autosome is activated during meiosis to replace that function, as is true for phosphoglycerate kinase-1 (27, 43), pyruvate dehydrogenase E1 $\alpha$  subunit (44), and glucose-6-phosphate dehydrogenase (45). Based on those studies, we propose that the  $\sim$ 70,000  $M_r$  CstF-64 is the product of an autosomal gene. In mouse, we have evidence for several autosomal homologs of *Cstf64*. We are exploring the possibility that one of these encodes the testis-specific CstF-64.

Curiously, the one other tissue in which we see low levels of  $\sim$ 70,000  $M_r$  CstF-64 expression is brain. Many testis-expressed genes also are found in brain (46–56). Similarly, there are many examples of differential mRNA processing in brain, the best known of which is the switch between calcitonin and calcitonin gene-related peptide (57, 58). Here, inclusion or exclusion of exon 4 depends on the interaction of CstF-64 with downstream sequences and other elements.

It is also unusual that expression of the somatic form of CstF-64 resumes after meiosis, suggesting that the somatic CstF-64 gene escapes inactivation after meiosis. Although most inactivated genes remain so after meiosis, a rarer class is reactivated in spermatids (59). In fact, our preliminary data suggest that the mRNA for the somatic CstF-64 is greatly overexpressed in postmeiotic germ cells (B.D., A.M.W., and C.C.M., unpublished work), reminiscent of the mRNA for the TATA-binding protein (60, 61). Several recent studies have highlighted the interdependence of polyadenylation and transcription (26, 62–65). It is therefore interesting to speculate that future data might uncover unique functional interactions of these essential processes in male germ cells.

**Note.** Beyer *et al.* (92) noted a 70,000  $M_r$  protein in purified CstF from calf thymus that is recognized by antibodies to CstF-64, which might be the same protein we describe here.

We thank André Reuss for excellent technical assistance, Gareth Howell for the gift of reagents, and Daniel Hardy and Simon Williams for critical reading of the manuscript. This work was supported by grants from the American Heart Association, Texas Affiliate and the Wendy Will Case Cancer Fund, Inc. (to C.C.M.), by the National Cancer Institute, Department of Health and Human Services, under contract with Advanced BioScience Laboratories (to N.A.J. and N.G.C.), and by a Helen Hedges Educational Charitable Trust Scholarship (to B.D.).

- Colgan, D. F. & Manley, J. L. (1997) *Genes Dev.* **11**, 2755–2766.
- Manley, J. L. (1995) *Curr. Opin. Genet. Dev.* **5**, 222–228.
- Wahle, E. (1995) *Biochim. Biophys. Acta* **1261**, 183–194.
- Keller, W. (1995) *Cell* **81**, 829–832.
- Proudfoot, N. J. & Brownlee, G. G. (1976) *Nature (London)* **263**, 211–214.
- Fitzgerald, M. & Shenk, T. (1981) *Cell* **24**, 251–260.
- Sheets, M. D., Ogg, S. C. & Wickens, M. (1990) *Nucleic Acids Res.* **18**, 5799–5805.
- Wilusz, J., Pettine, S. M. & Shenk, T. (1989) *Nucleic Acids Res.* **17**, 3899–3908.
- Conway, L. & Wickens, M. (1987) *EMBO J.* **6**, 4177–4184.
- Meijer, D., Hermans, A., von Lindern, M., van Agthoven, T., de Klein, A., Mackenbach, P., Grootegoed, A., Talarico, D., Valle, G. D. & Grosfeld, G. (1987) *EMBO J.* **6**, 4041–4048.
- Øyen, O., Myklebust, F., Scott, J. D., Cadd, G. G., McKnight, G. S., Hansson, V. & Jähnsen, T. (1990) *Biol. Reprod.* **43**, 46–54.
- Ravnik, S. E. & Wolgemuth, D. J. (1996) *Dev. Biol.* **173**, 69–78.
- Edwards-Gilbert, G., Veraldi, K. L. & Milcarek, C. (1997) *Nucleic Acids Res.* **25**, 2547–2561.
- Wilusz, J. & Shenk, T. (1988) *Cell* **52**, 221–228.
- Moore, C. L. & Sharp, P. A. (1985) *Cell* **41**, 845–855.
- Takagaki, Y., Manley, J. L., MacDonald, C. C., Wilusz, J. & Shenk, T. (1990) *Genes Dev.* **4**, 2112–2120.
- Takagaki, Y. & Manley, J. L. (1998) *Mol. Cell* **2**, 761–771.
- MacDonald, C. C., Wilusz, J. & Shenk, T. (1994) *Mol. Cell. Biol.* **14**, 6647–6654.
- Jenny, A. & Keller, W. (1995) *Nucleic Acids Res.* **23**, 2629–2635.
- Murthy, K. G. & Manley, J. L. (1995) *Genes Dev.* **9**, 2672–2683.
- Mann, K. P., Weiss, E. A. & Nevins, J. R. (1993) *Mol. Cell. Biol.* **13**, 2411–2419.
- Edwards-Gilbert, G. & Milcarek, C. (1995) *Mol. Cell. Biol.* **15**, 6420–6429.
- Takagaki, Y., Seipelt, R. L., Peterson, M. L. & Manley, J. L. (1996) *Cell* **87**, 941–952.
- Martincic, K., Campbell, R., Edwards-Gilbert, G., Souan, L., Lotze, M. T. & Milcarek, C. (1998) *Proc. Natl. Acad. Sci. USA* **95**, 11095–11100.
- Niwa, M., MacDonald, C. C. & Berget, S. M. (1992) *Nature (London)* **360**, 277–280.
- McCracken, S., Fong, N., Yankulov, K., Ballantyne, S., Pan, G., Greenblatt, J., Patterson, S. D., Wickens, M. & Bentley, D. L. (1997) *Nature (London)* **385**, 357–361.
- McCarrey, J. R. & Thomas, K. (1987) *Nature (London)* **326**, 501–504.
- Handel, M. A., Hunt, P. A., Kot, M. C., Park, C. & Shannon, M. (1991) *Ann. NY Acad. Sci.* **637**, 64–73.
- Dignam, J. D., Lebovitz, R. M. & Roeder, R. G. (1983) *Nucleic Acids Res.* **11**, 1475–1489.
- Gorski, K., Carneiro, M. & Schibler, U. (1986) *Cell* **47**, 767–776.
- Laemmli, U. K. (1970) *Nature (London)* **227**, 680–685.
- Towbin, H., Staehelin, T. & Gordon, J. (1979) *Proc. Natl. Acad. Sci. USA* **76**, 4350–4354.
- Ioannou, P. A., Amemiya, C. T., Barnes, J., Kroisel, P. M., Shizuya, H., Chen, C., Batzer, M. A. & de Jong, P. J. (1994) *Nat. Genet.* **6**, 84–89.
- Verma, R. S. & Babu, A. (1995) *Human Chromosomes: Principles and Techniques* (McGraw-Hill, New York).
- Copeland, N. G. & Jenkins, N. A. (1991) *Trends Genet.* **7**, 113–118.
- Jenkins, N. A., Copeland, N. G., Taylor, B. A. & Lee, B. K. (1982) *J. Virol.* **43**, 26–36.
- Rawlings, D. J., Saffran, C. D., Tsukada, S., Largaespada, D. A., Grimaldi, J. C., Cohen, L., Mohr, R. N., Bazan, J. F., Howard, M., Copeland, N. G., et al. (1993) *Science* **261**, 358–361.
- Russell, L. D., Ettlin, R. A., Sinha Hikim, A. P. & Clegg, E. D. (1990) *Histological and Histopathological Evaluation of the Testis* (Cache River Press, Clearwater, FL).
- Wingert, D., Reeves, R. & Magnuson, N. S. (1992) *Nucleic Acids Res.* **20**, 3183–3189.
- Foulkes, N. S., Schlotter, F., Pévet, P. & Sassone-Corsi, P. (1993) *Nature (London)* **362**, 264–267.
- Foulkes, N. S., Mellstrom, B., Benusiglio, E. & Sassone-Corsi, P. (1992) *Nature (London)* **355**, 80–84.
- Takagaki, Y., MacDonald, C. C., Shenk, T. & Manley, J. L. (1992) *Proc. Natl. Acad. Sci. USA* **89**, 1403–1407.
- McCarrey, J. R., Berg, W. M., Paragioudakis, S. J., Zhang, P. L., Dilworth, D. D., Arnold, B. L. & Rossi, J. J. (1992) *Dev. Biol.* **154**, 160–168.
- Dahl, H. H., Brown, R. M., Hutchison, W. M., Maragos, C. & Brown, G. K. (1990) *Genomics* **8**, 225–232.
- Hendriksen, P. J., Hoogerbrugge, J. W., Baarens, W. M., de Boer, P., Vreeburg, J. T., Vos, E. A., van der Lende, T. & Grootegoed, J. A. (1997) *Genomics* **41**, 350–359.
- Morales, C. R., Wu, X. Q. & Hecht, N. B. (1998) *Dev. Biol.* **201**, 113–123.
- Shackleford, G. M. & Varmus, H. E. (1987) *Cell* **50**, 89–95.
- Ayer-LeLievre, C., Olson, L., Ebendal, T., Hallbook, F. & Persson, H. (1988) *Proc. Natl. Acad. Sci. USA* **85**, 2628–2632.
- Law, H., Itkonen, O. & Lingwood, C. A. (1988) *J. Cell. Physiol.* **137**, 462–468.
- Persson, H., Pelto-Huikko, M., Metsis, M., Soder, O., Brene, S., Skog, S., Hofkelt, T. & Ritzen, E. M. (1990) *Mol. Cell. Biol.* **10**, 4701–4711.
- Tillakaratne, N. J., Erlander, M. G., Collard, M. W., Greif, K. F. & Tobin, A. J. (1992) *J. Neurochem.* **58**, 618–627.
- Komuro, I., Schalling, M., Jahn, L., Bodmer, R., Jenkins, N. A., Copeland, N. G. & Izumo, S. (1993) *EMBO J.* **12**, 1387–1401.
- Ewing, J. F. & Maines, M. D. (1995) *Endocrinology* **136**, 2294–2302.
- Schumacher, J. M., Lee, K., Edelhoff, S. & Braun, R. E. (1995) *J. Cell Biol.* **129**, 1023–1032.
- Srivastava, C. H., Monts, B. S., Rothrock, J. K., Peredo, M. J. & Pescovitz, O. H. (1995) *Endocrinology* **136**, 1502–1508.
- Defer, N., Marin, O., Poyard, M., Lienard, M. O., Jegou, B. & Hanoune, J. (1998) *FEBS Lett.* **424**, 216–220.
- Amara, S. G., Jonas, V., Rosenfeld, M. G., Ong, E. S. & Evans, R. M. (1982) *Nature (London)* **298**, 240–244.
- Lou, H. & Gagel, R. F. (1998) *J. Endocrinol.* **156**, 401–405.
- Hendriksen, P. J., Hoogerbrugge, J. W., Themmen, A. P., Koken, M. H., Hoeijmakers, J. H., Oostra, B. A., van der Lende, T. & Grootegoed, J. A. (1995) *Dev. Biol.* **170**, 730–733.
- Schmidt, E. E. & Schibler, U. (1997) *Dev. Biol.* **184**, 138–149.
- Schmidt, E. E. & Schibler, U. (1995) *Development (Cambridge, U.K.)* **121**, 2373–2383.
- Connelly, S. & Manley, J. L. (1988) *Genes Dev.* **2**, 440–452.
- Lanoix, J. & Acheson, N. H. (1988) *EMBO J.* **7**, 2515–2522.
- Dantonel, J. C., Murthy, K. G., Manley, J. L. & Tora, L. (1997) *Nature (London)* **389**, 399–402.
- Hirose, Y. & Manley, J. L. (1998) *Nature (London)* **395**, 93–96.
- Bunick, D., Johnson, P., Johnson, T. R. & Hecht, N. B. (1990) *Proc. Natl. Acad. Sci. USA* **87**, 891–895.
- Daneau, I., Houde, A., Ethier, J. F., Lussier, J. G. & Silversides, D. W. (1995) *Biol. Reprod.* **52**, 591–599.
- Korhonen, V. P., Halmekyto, M., Kauppinen, L., Myohanen, S., Wahlforss, J., Keinanen, T., Hyvonen, T., Alhonen, L., Eloranta, T. & Jänne, J. (1995) *DNA Cell Biol.* **14**, 841–847.
- Faber, P. W., King, A., van Rooij, H. C., Brinkmann, A. O., de Both, N. J. & Trapman, J. (1991) *Biochem. J.* **278**, 269–278.
- Ogawa, K. & Shimizu, T. (1993) *Biochim. Biophys. Acta* **1216**, 126–128.
- Furokawa, K. & Hotta, Y. (1993) *EMBO J.* **12**, 97–106.
- Watanabe, D., Yamada, K., Nishina, Y., Tajiama, Y., Koshimizu, U., Nagata, A. & Nishimura, Y. (1994) *J. Biol. Chem.* **269**, 7744–7749.
- Kong, M., Richardson, R. T., Widgren, E. E. & O'Rand, M. G. (1995) *Biol. Reprod.* **53**, 579–590.
- Kleene, K. & Flynn, J. F. (1987) *J. Biol. Chem.* **262**, 17272–17277.
- Miyasaka, H. & Li, S. S. (1992) *Biochem. Biophys. Res. Commun.* **185**, 818–825.
- Zhang, Y. L., Akmal, K. M., Tsuruta, J. K., Shang, Q., Hirose, T., Jetten, A. M., Kim, K. H. & O'Brien, D. A. (1998) *Mol. Reprod. Dev.* **50**, 93–102.
- Moss, S. B., Ferry, R. A. & Groudine, M. (1994) *Nucleic Acids Res.* **22**, 3160–3166.
- Noce, T., Fujiwara, Y., Ito, M., Takeuchi, T., Hashimoto, N., Yamanouchi, M., Higashinakagawa, T. & Fujimoto, H. (1993) *Dev. Biol.* **155**, 409–422.
- Hara, Y., Rovescalli, A. C., Kim, Y. & Nirenberg, M. (1992) *Proc. Natl. Acad. Sci. USA* **89**, 3280–3284.
- Denny, P. & Ashworth, A. (1991) *Gene* **106**, 221–227.
- Cunliffe, V., Koopman, P., McLaren, A. & Trowsdale, J. (1990) *EMBO J.* **9**, 197–205.
- Passananti, C., Corbi, N., Paggi, M. G., Russo, M. A., Perez, M., Cotelli, F., Stefanini, M. & Amati, P. (1995) *Cell Growth Differ.* **6**, 1037–1044.
- Muramatsu, T., Giri, P. R., Higuchi, S. & Kincaid, R. L. (1992) *Proc. Natl. Acad. Sci. USA* **89**, 529–533.
- Wolf, F. W., Sarma, V., Seldin, M., Drake, S., Suchard, S. J., Shao, H., O'Shea, K. S. & Dixit, V. M. (1994) *J. Biol. Chem.* **269**, 3633–3640.
- Bookbinder, L. H., Cheng, A. & Bleil, J. D. (1995) *Science* **269**, 86–89.
- Hardy, D. M. & Garbers, D. L. (1995) *J. Biol. Chem.* **270**, 26025–26028.
- Gao, Z. & Garbers, D. L. (1998) *J. Biol. Chem.* **273**, 3415–3421.
- Ueki, K. M. T. & Kincaid, R. L. (1992) *Biochem. Biophys. Res. Commun.* **187**, 537–543.
- Means, A. R., Cruzalegui, F., LeMagueresse, B., Needleman, D. S., Slaughter, G. R. & Ono, T. (1991) *Mol. Cell. Biol.* **11**, 3960–3971.
- Walter, L., Dirks, B., Rothermel, E., Heyens, M., Szpirer, C., Levan, G. & Gunther, E. (1994) *Mamm. Genome* **5**, 216–221.
- Russell, D. L. & Kim, K. H. (1996) *Biol. Reprod.* **55**, 11–18.
- Beyer, K., Dandekar, T. & Keller, W. (1997) *J. Biol. Chem.* **272**, 26769–26779.



# ANNUAL REVIEW OF BIOCHEMISTRY

VOLUME 71, 2002

## EDITORIAL COMMITTEE (2002)

TANIA A. BAKER  
ROBERT FILLINGAME  
MICHAEL GOTTESMAN  
DANIEL HERZUCH AG  
ROGER KORNBERG  
ROWENA G. MATTHEWS  
GREGORY PESKOV  
CHRISTIAN R. H. RAETZ  
CHARLES C. RICHARDSON  
JEREMY W. THORNER  
ROBERT TIAN

## RESPONSIBLE FOR THE ORGANIZATION OF VOLUME 71 (EDITORIAL COMMITTEE, 2000)

JOHN N. ABELSON  
MICHAEL GOTTESMAN  
DANIEL HERZCHLAG  
JUDITH P. KLINMAN  
MICHAEL E. O'DONNELL  
CHRISTIAN R. H. RAETZ  
CHARLES C. RICHARDSON  
DAVID W. RUSSELL  
ROBERT D. SIMONI  
THOMAS A. STEFZ  
JEREMY W. THORNER  
KEITH R. YAMAMOTO  
CHARLES CRAIK (GUEST)  
ROWENA G. MATTHEWS (GUEST)  
JOHN TANKUN (GUEST)

REFERENCE: C71  
SERIAL NUMBER: 10551300  
ATTORNEY DOCKET NUMBER: 10589-034-999

# EUKARYOTIC RIBONUCLEASE P: A Plurality of Ribonucleoprotein Enzymes

Shaohua Xiao,<sup>1</sup> Felicia Scott,<sup>1</sup> Carol A. Fierke,<sup>2</sup> and David R. Engelke<sup>1</sup>

<sup>1</sup>Department of Biological Chemistry, 3200 MSRB III, 1150 W. Medical Center Drive, University of Michigan, Ann Arbor, Michigan 48109-0606; e-mail: shxiao@umich.edu; phscott@umich.edu; engelke@umich.edu

<sup>2</sup>Department of Chemistry, University of Michigan, Ann Arbor, Michigan 48109-0606; e-mail: fierke@umich.edu

**Key Words** tRNA processing, RNase MRP, rRNA processing, nucleolus, ribozyme

**Abstract** Ribonuclease P (RNase P) is an essential endonuclease that acts early in the tRNA biogenesis pathway. This enzyme catalyzes cleavage of the leader sequence of precursor tRNAs (pre-tRNAs), generating the mature 5' end of tRNAs. RNase P activities have been identified in Bacteria, Archaea, and Eucarya, as well as organelles. Most forms of RNase P are ribonucleoproteins, i.e., they consist of an essential RNA subunit and protein subunits, although the composition of the enzyme in mitochondria and chloroplasts is still under debate. The recent purification of the eukaryotic nuclear RNase P has demonstrated a significantly larger protein content compared to the bacterial enzyme. Moreover, emerging evidence suggests that the eukaryotic RNase P has evolved into at least two related nuclear enzymes with distinct functions, RNase P and RNase MRP. Here we review current information on RNase P, with emphasis on the composition, structure, and functions of the eukaryotic nuclear holoenzyme, and its relationship with RNase MRP.

## CONTENTS

INTRODUCTION	166
Ribozyme Activity of RNase P RNA	166
Functions of RNase P in Bacteria and Eukaryotes	168
Precursor tRNA Substrate Recognition	168
Mechanism for the Pre-tRNA Cleavage Reaction	169
VARIOUS FORMS OF EUKARYOTIC RNASE P ACTIVITIES	170
Mitochondrial RNase P	171
Chloroplast RNase P	172
Eukaryotic Nuclear RNase P	172
RELATIONSHIP BETWEEN RNASE P AND RNASE MRP	180
SUBNUCLEAR LOCALIZATION OF RNASE P AND THE QUESTION OF COORDINATED RNA PATHWAYS	182
SUMMARY	184

0066-4154/02/0707-0165\$14.00

## Ribozyme Activity of RNase P RNA

The ribonucleoprotein nature of ribonuclease P (RNase P) has been proven for the holoenzymes in bacteria, archaea, and eukaryotic nuclei. Because the holoenzymes for RNase P and its closely related offshoot, RNase MRP, are highly complex nucleoproteins in eukaryotes (1–5), bacterial RNase P is composed of a large RNA subunit, usually ranging from 350 to 400 nucleotides, and one small protein subunit that contributes about 10% of the mass of the holoenzyme (reviewed in (6, 7); see also the RNase P database maintained by JW Brown at North Carolina State University, <http://www.mbio.ncsu.edu/RNaseP/main.html> (8)). All bacterial RNase P RNAs that have been tested are ribozymes, i.e., they can recognize and cleave substrates of precursor tRNA (pre-tRNA) in the absence of the protein subunit under high ionic strength *in vitro*. However, the small protein subunit binds near the catalytic core of RNase P RNA and interacts directly with the pre-tRNA substrates, thus facilitating pre-tRNA recognition and modulating RNase P RNA structure (9–13). The protein component is essential for RNase P function *in vivo*, but it is not known whether this absolute requirement is caused by these known interactions with substrates and RNase P RNA, or whether the protein plays additional roles in the cell.

In contrast to the situation in bacteria, eukaryotic RNase P RNA subunits by themselves have not yet been shown to cleave pre-tRNA substrates even though they contain the five most conserved “critical regions” that are postulated to form the catalytic center of the bacterial ribozymes (14, 15). One possible reason is an increased dependence on the more extensive protein content of the eukaryotic holoenzyme for correct three-dimensional folding of the RNA subunit (16). However, it is also possible that essential functions of substrate binding and catalysis have devolved to the protein subunits.

An intermediate level of protein dependence is found in archaea. RNA subunits from some *Methanobacteria*, *Thermococci*, *Pyrococci*, and halobacteria display traces of catalytic activity under extremely high salt concentrations in the absence of protein (17). Recently, it has been shown that the RNase P purified from *Methanohalobacter thermophilic* has at least four proteins that are apparent homologs of eukaryotic nuclear RNase P protein subunits (Table 1) (18). The RNase P RNA from this species is a ribozyme at very high salt levels, and the activity can be partially restored at lower salt levels by addition of the *Bacillus subtilis* RNase P protein (17). This last result is surprising, since neither the archaeal nor the eukaryotic proteins have any obvious sequence homology to the small bacterial protein subunits. This suggests that the RNase P protein can enhance catalytic activity either in a relatively nonspecific fashion or by interacting with the conserved catalytic core of RNase P RNA.

Yeast gene	Subunit composition of nuclear RNase P and RNase MRP from yeast ( <i>S. cerevisiae</i> ) and human cells	Human gene	Molecular weight (kDa)	Subunit type	MRP	RNase P	Archaeal RNase P	Reference
RP1	—	RNA	120	—	—	H1	—	—
NME1	—	RNA	112	—	—	—	—	—
POPI	POPI	Protein	100.5	9.84	hPOPI	hPOPI	80.86	168.169
POP3	POP3	Protein	22.6	9.57	—	—	81	73.79
POP4	POP4	Protein	32.9	9.26	RPP29 <sup>a</sup>	RPP29 <sup>a</sup>	18.82.88	16.83.89
POP5	POP5	Protein	19.6	7.79	hPOP5	hPOP5	MTTH11	1.18.87
POP6	POP6	Protein	18.2	9.28	—	—	—	—
POP7	POP7	Protein	15.5	9.34	RPP20	RPP20	1.85	—
POP8	POP8	Protein	15.5	4.57	—	—	—	—
RPP1	RPP1	Protein	32.2	9.76	RPP30	RPP30	MTTH688	18.83.85
RPP2	RPP2	Protein	16.3	9.99	RPP21	RPP21	MTTH1618	1.18.90
SNM1	SNM1	Protein	22.5	9.81	—	—	—	—
The relationship of these human proteins to the yeast protein subunits is currently unclear.								
RP07 is also known as RPP2 (84).								
RPP29 is also known as RPP4 (87).								

<sup>a</sup>Molecular weights are indicated only for the yeast RNA and protein subunits. PIs are presented only for the yeast protein subunits.

## Functions of RNase P in Bacteria and Eukaryotes

The best-defined function of RNase P is catalysis of the hydrolysis of a specific phosphodiester bond in pre-tRNAs, leaving a phosphate group at the 5' end of mature tRNA and a hydroxyl group at the 3' end of the leader. In bacteria, RNase P can also recognize and cleave non-tRNA natural substrates, including the precursors for 4.5S rRNA (19), 10S rRNA (a.k.a. tmRNA) (20), and the polycistronic *his* operon mRNA (21). Emerging studies of the bacterial holoenzyme have shown that the protein subunit interacts directly with the single-stranded leader of substrates, which enhances the affinity of pre-tRNA (22, 23) and allows RNase P to more efficiently cleave a variety of non-tRNA substrates (12, 24, 25).

In the case of eukaryotic RNase P, pre-tRNAs are the only natural substrates that have been identified to date. However, the possibility has not been ruled out that nuclear RNase P cleaves non-tRNA substrates. Indeed, several pieces of data have suggested that nuclear RNase P might have a range of substrates. For instance, a temperature-sensitive mutation in the RNA subunit of yeast RNase P RNA results in accumulation of an unusual 5.8S rRNA species *in vivo* (26), although this accumulation could well be due to indirect effects. Furthermore, purified nuclear RNase P from *Saccharomyces cerevisiae* cleaves deproteinized pre-tRNA at a large number of discrete sites (26) that are not recognized by the bacterial RNase P ribozyme (J.R. Chamberlain & D.R. Engelke, unpublished observations). In eukaryotes an additional enzyme, RNase MRP, has also evolved from RNase P (see below). This enzyme cleaves pre-tRNA, and possibly other substrates.

## Precursor tRNA Substrate Recognition

Since all pre-tRNAs in a cell appear to be processed by RNase P and sequences surrounding the cleavage site are not well conserved, RNase P must recognize the common structural elements of pre-tRNA substrates. Previous studies have shown that the major recognition determinants are localized to the mature and acceptor stems together form a minimal recognition determinant for the bacterial ribozyme (7). In contrast, efficient cleavage by eukaryotic RNase P requires both the TΨC/acceptor stems and at least two of the other three tRNA domains, i.e., D-, anticodon, and variable arms (27).

In addition to the tertiary structure of the mature tRNA, the 5' leader, 3' trailer, and internal loops can participate in binding to RNase P in bacteria and eukaryotes. In the ribozyme-catalyzed reaction, recognition is facilitated by interactions between a specific internal loop (P15) in the ribozyme and the 3'-CCA sequence that exists in most of the bacterial pre-tRNA transcripts (reviewed in 7, 27). However, this interaction appears to be less crucial for substrate recognition by the holoenzyme (28). For the holoenzyme, the protein subunit of *Bacillus subtilis* RNase P contains a single-stranded RNA-binding motif (29, 30) that directly contacts the substrate 5' leader and enhances the pre-tRNA affinity (22, 23, 31). Unlike the bacterial pre-tRNAs, eukaryotic pre-

tRNAs have a polyuridine [poly(U)] stretch as their 3'-trailing sequence and normally lack an encoded 3'-CCA. The nuclear RNase P RNA subunit also lacks the P15 internal loop that would base-pair with the substrate 3'-CCA in the bacterial enzyme. Interestingly, the 3' end in yeast pre-tRNA still seems to be recognized by the yeast nuclear RNase P. A tRNA containing single-stranded 3'-trailer sequence binds to the holoenzyme more strongly than does the mature tRNA (32).

There is some question as to whether recognition of pre-tRNA by the nuclear RNase P holoenzyme might be controlled *in vivo* by the structure of the 5' and 3' ends, or by possible association of the pre-tRNA with other proteins. The 5' leaders of eukaryotic pre-tRNAs are often purine-rich and can form base pairs to varying degrees with the U-rich 3' trailer, giving an extension of the amino acyl stem. When this base pairing in the amino acyl stem is strong in yeast pre-tRNAs, an unpaired bulge at the cleavage site is needed to enhance efficient cleavage by RNase P (33). However, it is not yet clear whether the leader/trailer stem extension must be fully unwound before cleavage by nuclear RNase P. Nor is it yet clear whether the pre-tRNA substrates are complexed with one or more additional proteins as they are cleaved by nuclear RNase P. Pre-tRNAs and other RNA polymerase III transcripts are thought to associate during transcription with the La protein antigen (34-36). In yeast it has been shown that nuclear pre-tRNAs are found associated with the La-like protein Lhp1p both before and after RNase P cleavage (37; S.L. Wolin, personal communications). Lhp1p, like human La antigen, binds primarily to the 3' poly(U) and would be expected to denature any structure between the 5'-leader and 3'-trailing sequences. Therefore, it is possible that inside the cell, La or other bound proteins serve as part of the pre-tRNA recognition mechanism by RNase P.

## Mechanism for the Pre-tRNA Cleavage Reaction

The pre-tRNA cleavage reaction catalyzed by RNase P has been studied in Bacteria, Archaea, and Eucarya; the best characterized of these is the bacterial ribozyme reaction. The most likely mechanism for the ribozyme reaction, which is proposed mainly by analogy with other, better studied reactions, is an S<sub>N</sub>2 attack of a solvent nucleophilic on the phosphate diester at the scissile bond with a trigonal bipyramidal transition state (38, 39).

The bacterial RNase P reaction consists of three basic steps: (a) pre-tRNA substrate binding; (b) scissile bond cleavage; and (c) product release. For bacterial RNase P, the affinity of RNase P for substrate and product is very dependent on the ionic conditions. However, at optimal salt conditions, the  $K_d$  of the interaction between the *B. subtilis* RNase P RNA alone and mature tRNA is about 3 nM. The addition of the protein component only modestly enhances the tRNA affinity but increases the pre-tRNA affinity 10<sup>4</sup> fold (22). Under many conditions, the steady-state kinetics for pre-tRNA cleavage, catalyzed by either RNase P RNA alone or the holoenzyme, are dominated by substrate binding and

product dissociation steps. At low substrate concentrations,  $k_{cat}/K_m$  often reflects the substrate association step, which is nearly diffusion controlled, whereas at saturating substrate concentrations, tRNA dissociation is the main rate-limiting step (22, 40). The rate constant of the cleavage step can be measured directly using transient kinetics or altered solution conditions (low pH or substitution of  $\text{Ca}^{2+}$  for  $\text{Mg}^{2+}$ ). The rate constant for the scissile bond cleavage is pH dependent, increasing at higher pH (7). At pH 7, the rate constant for the cleavage catalyzed by *B. subtilis* RNase P RNA is  $6 \text{ s}^{-1}$  (40), and this is only modestly enhanced (<10-fold) by addition of the protein subunit (22).

Studies of the eukaryotic RNase P reaction have been performed primarily with partially purified holoenzymes from nuclei and organelles (32, 41–44). Initial characterizations of the purified yeast nuclear RNase P reaction by steady-state analyses give a value for  $K_m$  of 9 nM, and a value for  $k_{cat}$  of  $1.3 \text{ s}^{-1}$  for cleavage of a specific yeast pre-tRNA $^{\text{Tr}}$  (32), although the kinetic parameters vary for different substrates. The second-order rate constant,  $k_{cat}/K_m$ , for the yeast nuclear holoenzyme is  $1.3 \times 10^8 \text{ M}^{-1} \text{ s}^{-1}$ . This result is close to the predicted rate constant for diffusion (32), which suggests that this is a very efficient enzyme.

Divalent metal ions are important cofactors for the RNase P reaction. The roles of divalent metal ions in RNase P RNA folding, substrate recognition, and catalysis of the ribozyme reaction have recently been reviewed (7, 45). Divalent metal ions enhance the affinity of *B. subtilis* RNase P RNA with tRNA more than 10<sup>3</sup>-fold (46), although cross-linking data suggest that RNase P RNA/tRNA association under some conditions may occur in the absence of metal ions (7). To date, all of the data indicate that divalent metal ions, preferably magnesium, are absolutely required for cleavage catalyzed by RNase P, although the exact functions of these metal ions have not yet been elucidated. Several possible metal-binding sites important for catalysis or substrate binding have been proposed in both the pre-tRNA substrate and the RNase P RNA by phosphorothioate modification experiments (reviewed in 45). In the case of yeast nuclear RNase P, studies of the utilization of metal ions by the holoenzyme reveal an even higher specificity for magnesium as the divalent metal ion, but the precise role of the divalent cations is currently unknown. Substitution of divalent cations in the holoenzyme has been difficult to document owing to irreversible inactivation of the enzyme by prolonged exposure to EDTA (47). Phosphorothioate substitutions at the cleavage site in pre-tRNA decrease the cleavage activity of yeast nuclear RNase P in a fashion similar to that seen with bacterial RNase P, which suggests a similar mechanism (41, 43).

## VARIOUS FORMS OF EUKARYOTIC RNASE P ACTIVITIES

In eukaryotes the RNase P functions have been partitioned into several different enzymes. The nuclear enzymic functions have been split into at least two distinct enzymes, RNase P and RNase MRP (see below). Pre-tRNA processing in

mitochondria and chloroplasts is carried out by activities that, in most studies, appear to be distinct from the nuclear holoenzymes. Characterization of RNase P-like activities from both mitochondria and chloroplasts has identified RNA-based enzymes to be similar to the nuclear and bacterial forms, as well as purely protein-based enzymes. The physiological relevance of these activities is currently under discussion (48–50).

### Mitochondrial RNase P

Mitochondrial RNase P activities have been characterized to various extents from yeasts, plants, vertebrates, and parasitic protozoa (7, 27, 51). The most detailed information on mitochondrial RNase P to date has been obtained by the studies of the budding yeast, *Saccharomyces cerevisiae*. The *S. cerevisiae* mitochondrial holoenzyme consists of an essential RNA subunit, Rpm1r (490 nucleotides, but varies in different strains) that is encoded in the mitochondrial genome, and a protein subunit, Rpm2p (105 kilodaltons (kDa)) that is encoded by a nuclear gene (52–55). Homologs of Rpm1r have also been identified in a number of yeasts with significant size variation compared to *S. cerevisiae* (56, 57), and a phylogenetic structure has been proposed for the RNA subunits that conforms to the bacterial consensus structure (58); the RNase P database, (8).

The Rpm2p protein subunit is required for mitochondrial RNase P activity, although like the nuclear holoenzyme proteins, it does not show significant sequence similarity to the bacterial RNase P proteins. Mutant alleles of Rpm2p cause accumulation of mitochondrial pre-tRNAs with the 5' end (54). In addition to pre-tRNA cleavage, Rpm2p is needed for the processing of the mitochondrial RNase P RNA subunit (59, 60). Furthermore, a complete deletion of the *RPM2* gene prevents fermentative growth of yeast cells (61), which suggests that Rpm2p has one or more unidentified functions that are essential for life.

The mitochondrial RNase P activities from other species are less well understood at this point. In the case of human mitochondrial RNase P, early reports identified an RNase P-like activity from HeLa cell mitochondrial fractions that appeared to be entirely protein based and to have a mitochondrial-specific substrate specificity (50, 62–64). On the other hand, an RNase P activity has recently been partially purified from a human HeLa cell mitochondrial fraction that contains the same nuclear-encoded H1 RNA as the human nuclear RNase P (65). Because this enzyme was identified by assays using heterologous pre-tRNA substrates of a type not cleaved by the protein-based activity, it remains possible that the H1-containing enzyme is a contaminant from a nonmitochondrial cellular compartment. It seems unlikely, though possible, that human mitochondria have two activities that cleave pre-tRNA. Because neither enzyme type has yet been shown to play a relevant role in mitochondrial pre-tRNA biosynthesis, the question of whether the mitochondrial RNase P activity requires an RNA subunit remains controversial.

## Chloroplast RNase P

The debate over RNase P RNA content in chloroplasts mirrors the discussion for the human mitochondrial enzymes. Chloroplast RNase P activities have been isolated from tobacco and spinach chloroplasts (7). Several lines of evidence suggest that spinach chloroplast RNase P might be a solely protein enzyme (reviewed in (7, 66)). (a) The buoyant density of the enzyme in CsCl is 1.28 g/ml, which is coincident with the total chloroplast protein; (b) spinach chloroplast RNase P activity is insensitive to micrococcal nuclease digestion; and (c) no RNA is detectable in the most purified enzyme preparations, even using sensitive assays. Moreover, spinach chloroplast activity seems to have a different cleavage mechanism than that of bacterial RNase P ribozyme (42), although the reaction is still very efficient. Spinach chloroplast RNase P binds to a substrate pre-tRNA with a  $K_m$  of 16 nM (42). Substitution of the pro-Rp nonbridging oxygen with sulfur at the scissile bond of pre-tRNA dramatically decreases the activity of bacterial ribozyme, but the activity of spinach chloroplast RNase P is hardly affected [reviewed in (7, 42, 67)]. A cautionary note concerning this RNase P activity: it has not yet been proven essential for pre-tRNA maturation *in vivo*.

In contrast, the ribonucleoprotein nature of RNase P has been functionally proven in the cyanelle of the primitive alga *Cyanophora paradoxa*; the cyanelle is a photosynthetic organelle derived from cyanobacteria and belongs to a different phylogenetic branch than the chloroplasts of green alga and higher plants (68). Although the RNA subunit alone is not catalytically active (69, 70), it is essential for RNase P activity, as indicated by the sensitivity of the holoenzyme to micrococcal nuclease treatment (27). The protein subunits of cyanelle RNase P have yet to be identified, but they constitute about 80% of the mass of the holoenzyme (68), an extensive protein content. Footprinting analysis of cyanelle RNase P RNA in the native holoenzyme also reveals extensive protection of the RNA by proteins, with a pattern similar to that of the yeast nuclear RNase P (68). In addition to cyanelle RNase P, sequences putatively like RNase P RNA have been identified in several chloroplast genomes, including those from maize (71), *Porphyra*, and *Nephroselmis olivacea* [the RNase P database (8)].

## Eukaryotic Nuclear RNase P

**COMPOSITION OF NUCLEAR RNASE P** Nuclear RNase P has been purified or partially purified from several eukaryotes, primarily from yeasts and vertebrates. The buoyant densities of eukaryotic holoenzymes are around 1.28–1.4 g/ml in  $\text{Cs}_2\text{SO}_4$ , which are less than those of the bacterial enzymes but between the densities of protein and RNA (7, 72). The protein complement contributes about 50% of the mass of the eukaryotic complex, as compared to 10% in the bacterial holoenzyme (73–77). These data, together with the ribonuclease sensitivity of eukaryotic RNase P activity, originally suggested the ribonucleoprotein nature of the nuclear holoenzymes.

In the yeast *S. cerevisiae*, the RNA subunit of nuclear RNase P, *RPR1* RNA, was first identified about a decade ago (78, 79). Genetic and immunoprecipitation studies identified the first four proteins associated with yeast nuclear RNase P: Pop1p, Pop3p, Pop4p, and Rpp1p (80–83). Biochemical purification (1) has confirmed the presence of the *RPR1* RNA, the above four proteins, and five additional polypeptides (Pop5p, Pop6p, Pop7p, Pop8p, and Rpp2p) (Table 1). Pop7p was also identified as a protein subunit because of homology to one of the human RNase P proteins (84). *In vivo* depletion studies have shown that all 10 subunits are essential for yeast viability and for RNase P activity (1, 79–83).

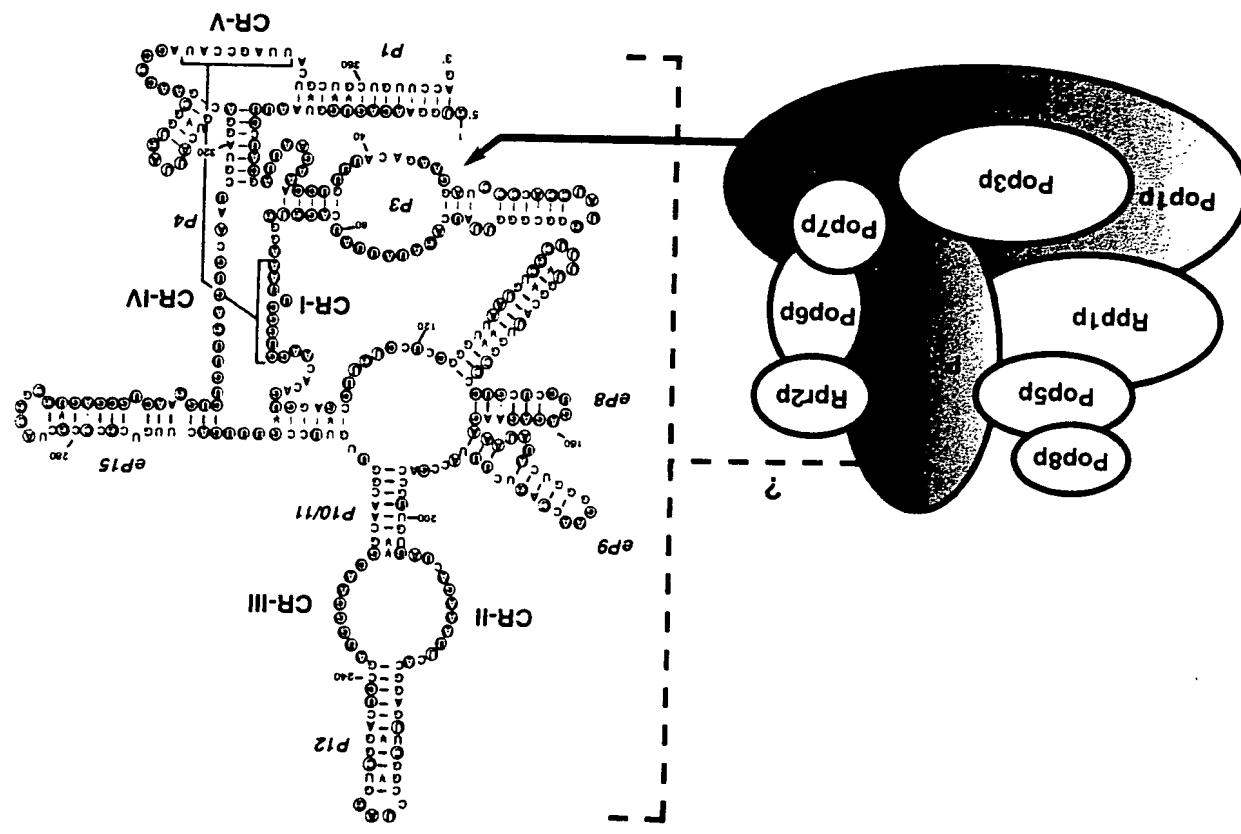
Purification of human RNase P from HeLa cells has revealed the existence of an RNA subunit, H1 RNA (73), and at least seven protein subunits that copurify with enzymatic activity, namely Rpp14, Rpp20, Rpp29, Rpp30, Rpp38, and Rpp40 (Table 1) (85). In addition, the human homolog of yeast Pop1p, hPop1, has been discovered independently by database searches, although the overall sequence similarity between the two proteins is low (86). Despite the lack of strong sequence similarity in *POP1* genes, antibody raised against hPop1 can immunoprecipitate the H1 RNA and RNase P activity from HeLa cell extracts (86). Thus, hPop1 is likely to be an integral component of human RNase P. In a separate study, Rpp29 has been identified as the homolog of yeast Pop4p and is also called hPop4 (87, 88). Database searches have recently identified additional human RNase P protein subunits based on sequence homology to other yeast nuclear RNase P proteins. hPop5 (the homolog of yeast Pop5p) and Rpp21 (the homolog of yeast Rpp2p) have proven to be integral subunits of human RNase P (89, 90). The relationship of other human RNase P proteins to yeast RNase P proteins remains unclear.

Recently, bioinformatic and biochemical studies have identified archaeal homologs of several yeast RNase P proteins, which suggests a eukaryotic-like protein content in archaeal RNase P. In *Methanothermophacter thermophilic*, the open reading frames MTH11 (POP4 homolog), MTH1687 (POPS homolog), MTH688 (RPP1 homolog), and MTH1618 (RPP2 homolog) have proven to be associated with RNase P activity (Table 1) (18).

**STRUCTURE AND FUNCTIONS OF THE NUCLEAR RNA SUBUNIT** The RNA component in eukaryotic nuclear RNase P has been isolated in several organisms (7, 8, 15). The corresponding genes appear to be transcribed by RNA polymerase III in yeast and vertebrates, but most of the information on the structure and function of the RNA subunit comes from the studies on the yeast enzyme. The *RPR1* RNA gene from *S. cerevisiae* is transcribed as a 486-nucleotide primary transcript (pre-*RPR1* RNA) (91). This large precursor, which might be a peculiarity of *Saccharomyces*, undergoes a single cleavage at the 5' end to remove the 84-nucleotide leader sequence and multiple cleavages to remove the 33-nucleotide 3'-trailing sequence to give the mature *RPR1* RNA (79). In normal *S. cerevisiae* strains, the ratio of the cellular level of pre-*RPR1* RNA and *RPR1* RNA is variable, but in actively growing cells it is about 1:9. Both pre-*RPR1*

RNA and *RPR1* RNA can be immunoprecipitated using affinity-tagged protein subunits (1, 80–83). Moreover, ribonucleoprotein complexes containing the pre-*RPR1* RNA and mature *RPR1* RNA have similar chromatographic characteristics during most of the biochemical purification steps (26). These data imply that the pre-*RPR1* RNA may be processed after the assembly into a ribonucleoprotein complex. Consistent with this, recent studies have shown that the pre-*RPR1* complex has catalytic activity (C. Srisawat, personal communication), which is not surprising considering that each of the protein subunits in the mature ribozyme can immunoprecipitate the pre-*RPR1* RNA. It is also possible that the pre-*RPR1* RNA has transiently associated proteins not found in the mature enzyme. For example, Lsm proteins have been found associated with the precursor, but not the mature, *RPR1* RNA (92).

A secondary structure for *RPR1* RNA has been proposed based on structure-sensitive RNA footprinting and phylogenetic studies (Figure 1) (15, 16, 93, 94). Comparison of the *RPR1* RNA structure with the bacterial consensus structure and other eukaryotic RNase P RNA structures has revealed both similarities and variability. A core structure is conserved among RNase P RNAs from bacteria and eukaryotes (14, 15, 95, 96). The conserved elements include five critical regions (CR-I through CR-V) carrying conserved nucleotides and several stems (P1, P2, P3, P4, P7, P10/11, and P12) at similar positions in different RNA structures (Figure 1). CR-I, CR-IV, and CR-V are also easily recognizable in the mitochondrial RNase P RNA from *S. cerevisiae*, but no obvious candidate sequence for CR-II or CR-III has been found (97). At this time no crystal structure for any RNase P RNA subunit has yet appeared, so information on structure and function of conserved regions comes mainly from mutagenesis, cross-linking, chemical probing, and computer modeling.



**Figure 1** Summary of the yeast two-hybrid and three-hybrid test results of yeast (*S. cerevisiae*) nuclear RNase P. Ovals represent the protein subunits of yeast nuclear RNase P (Table 1). Overlaps of the ovals indicate a positive result in the two-hybrid test (14a). The shaded ovals of Pop1p and Pop4p indicate that they are the only protein subunits that interact specifically with the *RPR1* RNA in a three-hybrid test. One of the recognition sites in the *RPR1* RNA for Pop1p binding is the P3 loop, as indicated by the arrow (116). The binding sites for Pop4p are currently unknown, which is represented by the dotted bracket and a question mark. The predicted secondary structure of *RPR1* RNA is adapted from the model proposed by Frank et al. (15). Five critical regions, CR-I to CR-V, are numbered based on the Chen & Pace nomenclature (14). P represents helical regions, with numbers assigned according to the bacterial structure (101). ep indicates the eukaryotic-paired regions whose homology to particular bacterial structures is uncertain but which occupy the same positions as in the bacterial structure (15). Nucleotides in filled circles show protection from chemical modification and nucleic acid attack in the holoenzyme, whereas the nucleotides in open circles indicate exposure to solution (16).

The P4 helix, which is formed by the base pairing of CR-I and CR-V, is postulated to be the catalytic center of the bacterial enzymes by site-directed mutagenesis, phosphorothioate modification-interference studies, and deletion analysis (96, 98–103). It is also suggested to be a potential binding site for the bacterial RNase P protein subunit (13; S. Niranjanakumari & C.A. Fierke, personal communication). For yeast nuclear RNase P, mutagenesis studies of the conserved nucleotides within and near P4 have suggested an important role in pre-tRNA binding, catalysis, and *RPR1* RNA maturation (103). Several substitutions of the conserved nucleotides cause more than a 10-fold reduction in the catalytic efficiency ( $k_{cat}/K_m$ ) of the holoenzyme and/or a severe decrease in the cellular level of the mature *RPR1* RNA, which is thought to be caused by an *RPR1* RNA maturation defect (103). Previous RNA footprinting studies have shown that P4 helix is protected from nuclease attacks in the RNase P holoenzyme (16), which suggests that it is not on the surface of the holoenzyme.

Function and accessibility to chemical reagents is magnesium dependent in yeast RNase P (104, 105). Deletional studies of the yeast RNase P RNA have shown that the distal P12 stem is dispensable in vivo, although the internal loop containing the CR-II and CR-III regions is essential for yeast viability and RNase P activity (106). The CR-II region contains a consensus sequence AGARA (R = purine), which is conserved in all yeast RNase P RNAs and similar to the bacterial CR-II consensus (15, 104). Mutational studies of the AGAA A sequence in CR-II of the *S. cerevisiae* RNase P RNA indicate a function in magnesium utilization as measured by an increase in optimum  $Mg^{2+}$  concentration for the mutants (104). In addition, steady-state kinetic studies of the mutant holoenzymes show effects primarily on turnover ( $k_{cat}$ ), with modest changes in  $K_m$  (104). In contrast, it has been shown that the analogous CR-II/III region in the bacterial ribozyme is not absolutely essential for cleavage, rather playing a role in substrate discrimination (107–111). Previous studies on bacterial RNase P RNA have shown that substrate binding requires divalent cation-mediated interactions with the CR-II and CR-III conserved loop and that the P10/11-P12 domain is subject to lead-induced hydrolysis and  $Mg^{2+}$  cleavage at high pH (98, 112). Cross-linking studies on the *Escherichia coli* and *B. subtilis* ribozymes also place CR-II and CR-III in proximity to the substrate aminoacyl stem (J.M. Nolan, personal communication).

Extensive mutagenesis of the CR-IV region of the *RPR1* RNA gives large reduction in  $k_{cat}$ , with little effect on  $K_m$  (103). In addition, the maturation of *RPR1* RNA does not seem to be affected in the strains containing the RNA mutated in this region (103). Studies on the bacterial enzyme have demonstrated the importance of this region in RNA binding (28, 113, 114). Thus, CR-IV might also be located in the proximity of the active site of the eukaryotic holoenzyme.

Structural elements distinct from the bacterial RNase P RNA subunits also exist in the nuclear RNAs; the most obvious are the P3 and P15 regions. Although the RNA species from both kingdoms contain a P3 stem, the eukaryotic

P3 has a more complex helix-loop-helix structure (Figure 1) (14, 15, 93, 95). P3 has been proposed as a protein binding site in bacterial RNase P holoenzyme, as indicated by deletional analysis and protein-directed RNA cleavage (13, 115; S. Niranjanakumari & C.A. Fierke, personal communication). In yeast RNase P, a mutated P3 loop causes defects in pre-tRNA processing and *RPR1* RNA maturation in vivo (116). Changes of the conserved nucleotides in the P3 loop also disrupt the specific interaction between *RPR1* RNA and one of the protein subunits (Pop1p) (116). This may cause a defect in the assembly of a ribonucleoprotein complex *in vivo*, which in turn may cause the loss of maturation of the RNA subunit. The P3 region in human RNase P RNA is also a polypeptide-binding site (117, 118), which suggests that both the bacterial and eukaryotic P3 regions are important for the assembly of RNase P holoenzyme. In addition to (or as a result of) its role in assembly, the eukaryotic P3 domain appears to be required for the proper subcellular localization of the RNase P RNA. In human RNase P, deletion of the P3 region causes the RNA component to fail to localize to the nucleolus (119).

In contrast to P3, the P15 region is often more complex in bacteria than in yeast and completely disappears in human RNase P RNA. Because it is not clear whether the eukaryotic structure in this region plays a role similar to the bacterial P15, it is termed ep15 (Figure 1) (15). The function of the P15 region is better understood in the bacterial RNase P ribozyme reaction than that of the holoenzyme. In the ribozyme reaction, the P15 loop has been shown to base pair with the 3'-CCA sequence of the substrate pre-tRNA (and product) (28, 120). Kinetic and biophysical analysis suggested that there might be a metal ion-binding site in the bacterial P15 loop (121–123). However, this interaction may be less important in the cleavage catalyzed by the bacterial holoenzyme (123). The function of the ep15 loop, if any, is not known.

A consensus secondary structure of eukaryotic RNase P RNA has recently been proposed (15). In addition to the conserved and variant regions mentioned above, there are elements in the eukaryotic consensus structure that might have counterparts in the bacterial consensus structure. Helices ep8 and ep9 (Figure 1) could be the homologs of the corresponding helices in bacterial RNase P RNA, although the functional equivalence requires further evidence. The fungal ep8 stem-loop structures have the NUCAG sequence (N = A, U, C, or G), whereas most of the ep9 hairpins contain GNR (R = purine) tetraloops (15). Studies of hairpins from ribosomal RNA and other sources have suggested that a tetraloop structure could increase the thermodynamic stability of an RNA duplex, or act as a docking site for intramolecular or intermolecular RNA-RNA interactions (124–130). In *E. coli* RNase P RNA (M1 RNA), a mutation in the L9 loop results in a reduction in  $K_m$  and  $k_{cat}$  of the holoenzyme-catalyzed reaction with pre-4.5S RNA as substrate (131), but little effect on pre-tRNA processing *in vitro*. The presence of the conserved catalytic core structure of bacterial RNase P RNA (CR-I through V) in the eukaryotic nuclear RNA subunit strongly suggests that the nuclear RNA is serving functions similar to the bacterial counterpart,

despite the presence of a much larger protein contingent. The importance of the critical regions in the yeast *RPR1* RNA for RNase P activity supports the hypothesis that nuclear RNase P is an RNA-based enzyme.

**FUNCTIONS OF THE PROTEIN SUBUNITS** The protein complement of eukaryotic RNase P is much more complex than that of the bacterial holoenzyme. Bacterial RNase P has a small, highly basic protein subunit, with a molecular mass of 12–14 kDa. In contrast, RNA footprinting of the yeast holoenzyme has shown that most of the RNA subunit is covered by proteins (Figure 1) (16). The purified nuclear RNase P from *S. cerevisiae* has 9 tightly associated, essential protein subunits, with molecular masses ranging from 15.5 to 100.5 kDa (Table 1). Except for Pop8p and Pop5p, the yeast nuclear RNase P protein components are very basic, with pI values higher than 9. The pI values of Pop8p and Pop5p are 4.6 and 7.8, respectively. Sequence analysis has revealed several putative signals for nuclear localization in some of the yeast protein subunits, but defined RNA-binding motifs are lacking, other than clusters of basic residues. A motif containing repetitive KKD/E sequences is found in most of the protein subunits (1). The functions of the basic clusters, including the KKD/E motif, are not clear, and it is plausible that these might participate in RNA-protein interaction, protein-protein interaction, and/or nuclear localization.

In fact, very little is known about any of the functions of individual protein subunits of nuclear RNase P, although several of the proteins expressed in bacteria exhibit nonspecific RNA-binding capability *in vitro* (F. Houser-Scott, C.E. Millikin & D.R. Engelke, unpublished observations). Depletion of most of the yeast protein subunits results in a drastic reduction in the cellular level of mature *RPR1* RNA (1, 82–84). This might reflect incorrect assembly of the yeast holoenzyme, so that pre-*RPR1* RNA cannot be processed, or destabilization of the mature *RPR1* RNA. The sole exception to this appears to be Pop3p, which is essential but does not dramatically affect the ratio of precursor to mature *RPR1* RNA (81). Recent studies on Pop3p have revealed that it can bind to *RPR1* RNA, pre-tRNA, and single-stranded RNAs (132). At this point, it is not clear whether the proteins bind to the RNA subunit as a preformed complex or are added to the RNA one at a time.

In the case of human RNase P protein subunits, a recent study has shown that the recombinant Rpp20 displays ATPase activity (133). However, the ATPase activities of Rpp20 and RNase P holoenzyme do not have the same optimal reaction conditions (133). The generality of this ATPase activity is unclear at present since the apparent yeast homolog of Rpp20, Pop7p, does not have the ATPase signature motif. Thus, the *in vivo* ATPase activity of Rpp20 and its importance to RNase P function await further analysis. In addition, a two-hybrid screen suggested that Rpp20 interacts with Hsp27, a small heat shock protein 27 (134). The function of Hsp27 in RNase P activity might be regulatory, since the addition of Hsp27 stimulates RNase P activity in a concentration-dependent manner (134).

It is likely that one or more protein subunits of eukaryotic RNase P mimic the functions of the bacterial protein in enhancing substrate binding and pre-tRNA discrimination (23). In human RNase P, Rpp21 and Rpp14 might interact directly with substrates, since they have been shown to bind to pre-tRNA in a gel mobility shift assay (90). However, the specificity for the sites of binding on the tRNA have not yet been defined. A labeled pre-tRNA substrate has been reported to form cross-links with protein components from the *Tetrahymena thermophila* and partially purified human RNase P (135). The yeast nuclear holoenzyme appears to have additional binding sites for single-stranded RNAs. Unlike the bacterial enzyme, it strongly binds and is inhibited by single-stranded homopolymer [poly(G) and poly(U) > poly(A) > poly(C)] (32). One role for such single-stranded-binding sites in the proteins might be recognition of the 3' trailer or the 5' leader of pre-tRNAs (32). A much wider role for the protein components in substrate discrimination still awaits discovery. Although the full range of physiological substrates of RNase P has not been tested, preliminary studies suggest that the nuclear holoenzyme is relatively promiscuous in cleaving naked RNA (26). Since nuclear RNAs are expected to be complexed with proteins in most cases *in vivo*, the protein components of nuclear RNase P could help discriminate among potential ribonucleoprotein substrates rather than simply recognizing RNA determinants. Indeed, part of the reason for the increased protein complexity of the nuclear enzyme could be to specify which RNAs are not to be cleaved.

The additional eukaryotic protein subunits do not appear to substantially alter the chemistry of the cleavage reactions. Studies on cleavage of phosphorothioate modified substrates have suggested that the protein-based spinach chloroplast activity uses different chemistry than the RNase P ribozyme (42) but that nuclear holoenzymes mirror the RNA-based cleavage (41, 43). It is also possible that entirely distinct enzymatic activities, such as RNA modification enzymes, might be coordinated with RNase P cleavage by being part of the same complex, or recruited to the complex through protein-protein interaction. Previous work has linked a tRNA methyltransferase activity to the stability of *RPR1* RNA (136), although none of the identified subunits have any detectable sequence homology to known tRNA processing or modification enzymes.

The subnuclear localization of several pre-tRNA processing enzymes [137–139; see below] has led to the hypothesis that the nuclear pre-tRNA pathway is spatially organized. It is entirely possible that one or more protein subunits are important for localization of RNase P within the nucleus. This might be required for the spatial organization of pre-tRNA biosynthesis, the maturation of RNase P holoenzyme, or participation of the enzyme in other undiscovered processing pathways.

**ARCHITECTURE OF THE EUKARYOTIC NUCLEAR HOLOENZYME** Although the crystal structure of eukaryotic RNase P holoenzyme will be difficult to obtain because of its low natural abundance and subunit complexity, a preliminary map of how the subunits fit together has been proposed by yeast two-hybrid and

three-hybrid tests (140, 141) and by cross-linking studies. The presence of both RNA and protein components in RNase P suggests that some proteins may make direct contacts with the RNA moiety and other proteins might be recruited through protein-protein interactions.

In the case of the *S. cerevisiae* nuclear RNase P, the yeast two-hybrid assay has revealed that all the protein components interact with at least one other protein subunit (Figure 1) (141a). Pop1p, Pop4p, Pop1p, Pop6p, Pop6p, and Rpr2p are involved in multiple protein-protein interactions; Pop4p and Rpr2p are able to self-associate. By the use of the yeast three-hybrid system, the RNA-protein interactions within yeast nuclear RNase P have been at least partially delineated (Figure 1). Pop1p and Pop4p are the only protein subunits that specifically bind to the *RPR1* RNA in the three-hybrid test (141a). A Pop1p binding site has been mapped to the P3 region of *RPR1* RNA (116). The RNA binding site(s) for Pop4p remain to be discovered. It is worth noting that other protein subunits might make direct contacts with the RNA, but the interactions could be dependent on the formation of a multiprotein complex.

Using a similar genetic analysis and cross-linking experiments, the subunit interaction map of human nuclear RNase P has been examined. Eight of the ten known protein subunits of human nuclear RNase P were tested in a yeast two-hybrid assay for their abilities to interact with each other (134). Rpp21 (human Rpr2p homolog), Rpp29/hPop4, Rpp30 (human Rpp1p homolog), Rpp38, Rpp40, and hPop1 are involved in extensive, but weak, protein-protein interactions among the RNase P subunits. Rpp21, Rpp29, Rpp30, and Rpp38 bind to the H1 RNA, as indicated by both a yeast three-hybrid test and UV cross-linking (142), although the RNA recognition elements for these proteins are currently unknown. Furthermore, the sera from some patients with autoimmune diseases react with a 40-kDa polypeptide, designated as To antigen, within human RNase P holoenzyme. The To antigen can be cross-linked to nucleotides 20–75 of human RNase P RNA (the To binding site), which encompasses the P3 domain of the RNA component (118). However, the relationship between the To/Th antigen and the known protein subunits of RNase P is unclear at present. The protein-protein interactions and protein-RNA interactions in the yeast nuclear RNase P are not identical to those observed in the human holoenzyme, although there is significant similarity. Considering the low identity between the protein homologs and the variation in the RNA structures, it is possible that the subunit interaction surfaces are different in these organisms, resulting in subtle changes in the strength of binary subunit contacts and the spatial organization of RNase P holoenzymes.

RNA primers for DNA replication in vitro (143, 144). However, most of the RNase MRP RNA is observed in the nucleus (145–148), where the enzyme plays an important role in pre-rRNA processing (3, 149, 150). In yeast, RNase MRP has been shown to cleave pre-rRNA at site A3 within the first internal transcribed spacer (ITS 1), which is essential for generating the mature 5.8S rRNA in vivo (80, 151–153). Both the RNA and protein subunits of RNase MRP are essential for yeast viability, even though neither the cleavage of pre-rRNA at site A3 nor the mitochondrial DNA replication are required for life (149). Therefore, RNase MRP may have additional essential functions that have not yet been discovered. Indeed, a yeast strain carrying a temperature-sensitive mutant of RNase MRP RNA displays a defect in pre-rRNA processing at sites other than A3 at nonpermissive temperature (149). However, the essential role of RNase MRP might not involve the pre-rRNA pathway. Mutagenesis of yeast RNase MRP has shown that the enzyme plays a role in plasmid segregation (154), and it is entirely possible that RNase MRP has additional substrates that have not yet been identified.

Characterization of RNase MRP primarily comes from yeast and human cells. Yeast nucleolar RNase MRP has an RNA component, *NME1* RNA, and at least nine protein subunits identified by genetics and immunoprecipitation (Table 1). Eight of these nine proteins are also tightly associated with the purified yeast nuclear RNase P. This protein subunit overlap and the similarity of the RNA subunits to the RNase P RNAs (155; see below) strongly indicate that RNases P and MRP are closely related descendants of an ancestral eukaryotic enzyme in which the RNA subunit has diverged and unique protein subunits have evolved accordingly.

The one known unique protein subunit of yeast RNase MRP, Snm1p, is clearly distinct from the unique protein subunit of RNase P, Rpr2p, yet the two proteins have some properties in common. In two-hybrid tests, both proteins self-associate and bind to Pop4p (141a). One would predict that the Snm1p subunit might perform functions that set MRP aside from RNase P, such as substrate discrimination or enzyme localization. Indeed, a putative zinc-cluster domain has been identified in Snm1p, and the protein can bind to *NME1* RNA in gel mobility shift and Northwestern assays (156). In contrast, Rpr2p has not been found to bind the *RPR1* RNA (141a). It is worth noting that since RNase MRP has not yet been purified from any source, the holoenzyme could contain additional subunits not found in RNase P.

Secondary structures of RNase MRP RNA have been proposed based on phylogenetic studies (157–160). RNase MRP RNA can fold into a cage-shaped secondary structure similar to that of the RNase P RNA. RNase MRP has at least three of the critical regions (CR-I, CR-IV, and CR-V) that are conserved in all RNase P RNAs, and it also shares the eukaryotic P3 loop. It is not known at present whether the equivalents of CR-II and CR-III are present in modified form in RNase MRP RNA.

## RELATIONSHIP BETWEEN RNASE P AND RNASE MRP

Nuclear RNase P is closely related to another ribonucleoprotein complex, RNase MRP, which has been found only in eukaryotes. RNase MRP was originally identified as an RNA-containing endoribonuclease that cleaves the mitochondrial

The P3 internal loop sequences of yeast RNase P RNA and RNase MRP RNA appear to be functionally interchangeable (161). The P3 region is thought to be a specific binding site for a common protein of RNase P and RNase MRP, although different types of experiments done with yeast and human enzymes have identified different probable binding partners. In a yeast three-hybrid test, Pop1p binds to the P3 region, alone or in the context of *RPR1* RNA, in a sequence-specific fashion (116). In human nuclear RNase P and RNase MRP, several polypeptides have been shown to require the P3 domain for binding or bind to a region close to P3 (117, 118, 162).

Information concerning the arrangement of proteins on the RNA subunit gives different answers for yeast versus human RNase MRP. Although the yeast P3 domain (equivalent to the Th/To binding-site in human RNase MRP RNA) binds directly to Pop1p, the human Pop1 does not cross-link to MRP RNA (162). However, the association of hPop1 with the hPop1 with the To/Th binding site in the RNA (162). Instead of hPop1, cross-linking experiments have shown that three polypeptides, with molecular masses of 20, 25, and 40 kDa, are in close contact with the To/Th-binding site (nucleotides 15–87) of the human RNase MRP RNA (117, 162). A 40-kDa protein is also cross-linked to the To/Th region (nucleotides 20–75) of human RNase P RNA (118). However, it is not known whether the 40-kDa polypeptides in these studies are identical or whether they correspond to any of the identified yeast RNase P/MRP subunits. Another 40-kDa protein, which is identical to Rpp38, has been found to bind to an RNA site that is on the 3' side of the To/Th-binding site (162).

## SUBNUCLEAR LOCALIZATION OF RNASE P AND THE QUESTION OF COORDINATED RNA PATHWAYS

In bacteria, RNase P is involved in the processing of tRNAs, rRNAs, and in specialized cases, mRNAs. The existence of RNAs within the large, ribosomal and rRNA pathways, since tRNA processing by RNase P is occurring as part of rRNA processing. In eukaryotes tRNA sequences are no longer found in the pre-rRNA transcription units, and cleavage between the large rRNA subunits is accomplished by a derivative of RNase P, RNase MRP, that has adapted to this task in the absence of a tRNA substrate structure. Given that nuclear tRNA and rRNA cleavage is accomplished by distinct enzymes, the question persists: Is there still any physical linkage between the two pathways?

In eukaryotes, most ribosome biogenesis, including the processing of pre-rRNA by RNase MRP, takes place in dense, subnuclear regions called nucleoli. Pre-rRNA biosynthesis had been assumed to be distributed throughout the nucleoplasm, but emerging data on the localization of nuclear RNase P suggest this might not be the case. In the yeast *S. cerevisiae*, the RNase P cleavage is an early step in the nuclear rRNA maturation pathway. Most nuclear pre-rRNAs and

both the precursor and mature RNase P RNAs are localized primarily to the nucleolus, with additional foci in the nucleoplasm (139, 161; E. Bertrand & D.R. Engelke, unpublished observations). Although one long-lived pre-tRNA<sup>le</sup> (164), has been found throughout the nucleus (138), the predominantly nucleolar localization results clearly suggest that the early steps in nuclear tRNA processing pathway are compartmentalized, rather than distributed evenly in the nucleoplasm. The partial colocalization of pre-tRNA and pre-rRNA maturation pathways in the nucleolus might reflect the possible interplay between tRNA and rRNA biogenesis. Considering that RNase P and RNase MRP share eight out of nine protein subunits, it is not surprising that they might have at least partially common subcellular localization.

In mammalian cells, *in situ* hybridization shows that endogenous RNase MRP RNA is predominantly nucleolar but that RNase P-H1 RNA is localized throughout the nucleus (119). Microinjection of labeled RNase P RNA demonstrated that the RNA is first localized in the dense fibrillar component of the nucleolus and subsequently redistributed throughout the nucleus (119). This would be consistent with some aspect of RNase P assembly taking place in the nucleolus and would also be consistent with the observation that precursors to yeast *RPR1* RNA are nucleolar (E. Bertrand & D.R. Engelke, unpublished observations). The To/Th antigen-binding site (nucleotides 25–75 in H1 RNA) is required for the initial nucleolar localization of the RNA, and the more 3' region of the RNA (nucleotides 89–341 in H1 RNA) is necessary for its nucleoplasmic redistribution (119). The location of RNase P in eukaryotic nuclei is somewhat controversial, however, since earlier reports suggested that concentrated foci of RNase P RNA were localized on the fringes of some HeLa cell nucleoli, termed the perinucleolar compartments (PNCs) (165).

In addition to detection of the RNA subunit, subcellular localization of the human RNase P protein components has also been studied. These experiments generally give answers that are difficult to interpret. Since many of the subunits are common to both RNases P and MRP, it is never clear whether the protein being visualized is associated with one of the enzymes, with both, or with neither. Using indirect immunostaining, most of the endogenous hPop1 is found in the fibrillar compartment of the nucleolus, although a weak, homogeneous staining in the nucleoplasm is also observed (86). Rpp29, Rpp38, Rpp14, and hPop5 are mainly detected in the nucleolus (89, 166). Rpp29 and Rpp38 have also been located in Cajal bodies (166). However, localization should not necessarily be considered static, since Rpp29 has been shown to rapidly shuttle between nucleolus and nucleoplasm in human cells (167).

Rpp21, the human homolog of the yeast RNase P-specific protein subunit Rpp2p, is primarily nucleoplasmic with a small amount in the nucleoli (90). Whether Rpp21 is associated only with RNase P activity in human cells has yet to be confirmed. An alternatively spliced form of human Rpp21, which does not associate with RNase P RNA, is detected primarily in the nucleoli (90). The significance of this alternatively spliced Rpp21 form is not immediately clear, but

its existence serves to emphasize how little we know about the functions of these proteins. For the yeast subunits, we know only that they are essential for life, for RNase P activity, and in most cases, for RNase MRP activity. These proteins could also play additional, completely unknown roles in the cell.

## SUMMARY

Ribonuclease P is one of two ancient ribonucleoprotein enzymes that are found in all living organisms (the other is the ribosome). The existence of a ribozyme RNA subunit in the bacterial enzymes initially caused much excitement, contributing to the creation of an RNA World hypothesis in which RNAs were the primordial macromolecules that both carried genetic information and catalyzed reactions. In eukaryotic nuclei the RNA subunit still has most of the features that are conserved in the bacterial ribozymes, but the holoenzyme has gained nearly 20 times as much protein and has simultaneously lost the ability to bind and cut pre-tRNA substrates without the protein complement. Although the pre-tRNA cleavage mechanism and functions of the nuclear RNase P RNA subunit are thought to be similar or identical to the bacterial counterpart, the reasons for the huge increase in protein content of eukaryotic holoenzymes are largely mysterious. The nuclear enzyme clearly has to deal with issues that are not likely to be seen by the bacterial enzymes—subcellular localization during biogenesis, possible spatial organization of the pre-tRNA pathway, and a potential need to recognize more complex substrates that ribonucleoprotein contacts.

## ACKNOWLEDGMENTS

We are grateful to T.A. Hall, J.W. Brown, S.L. Wolin, S. Nirajananakumari, W. Rossmanith, and J.M. Nolan for communicating results prior to publication, and to N.C. Martin and J.W. Brown in particular for helpful suggestions on the manuscript. This work is supported by the National Institutes of Health grants GM 34869 (to DRE) and GM 55387 (to CAF), the Anthony and Lilian Lu Fellowship (to SX), and postdoctoral fellowships from the UNCF-Pfizer and the Ford Foundation (to FHS).

The *Annual Review of Biochemistry* is online at <http://biochem.annualreviews.org>

## LITERATURE CITED

- Chamberlain JR, Lee Y, Lane WS, Engelke DR. 1998. *Genes Dev.* 12:1678–90.
- Reddy R, Shumba S. 1995. *Mol. Biol. Rep.* 22:81–85.
- Xiao SH, Houser-Scott F, Engelke DR. 2001. *J. Cell. Physiol.* 187:11–21.
- Pace NR, Brown JW. 1995. *J. Bacteriol.* 177:1919–28.
- Frank DN, Pace NR. 1998. *Annu. Rev. Biochem.* 67:153–80.
- Brown JW. 1999. *Nucleic Acids Res.* 27:314.
- Nirajananakumari S, Kurz JC, Fierke CA. 1998. *Nucleic Acids Res.* 26:3090–96.
- Vioque A, Arnez J, Altman S. 1988. *J. Mol. Biol.* 202:835–48.
- Talbot SJ, Altman S. 1994. *Biochemistry* 33:1399–405.
- Loria A, Nirajananakumari S, Fierke CA, Pan T. 1998. *Biochemistry* 37:15466–73.
- Biswas R, Ledman DW, Fox RO, Altman S, Gopalan V. 2000. *J. Mol. Biol.* 296:19–31.
- Chen JL, Pace NR. 1997. *RNA* 3:557–60.
- Frank DN, Adamidi C, Ehringer MA, Pitulle C, Pace NR. 2000. *RNA* 6:1895–904.
- Tranguch AJ, Kindelberger DW, Rohlman CE, Lee JY, Engelke DR. 1994. *Biochemistry* 33:1778–87.
- Pannucci JA, Haas ES, Hall TA, Harris JK, Brown JW. 1999. *Proc. Natl. Acad. Sci. USA* 96:7803–8.
- Hall TA, Brown JW. 2002. *RNA*. In press.
- Bothwell ALM, Garber RL, Altman S. 1976. *J. Biol. Chem.* 251:7709–16.
- Komine Y, Kitabayake M, Yokogawa T, Nishikawa K, Inokuchi H. 1994. *Proc. Natl. Acad. Sci. USA* 91:9223–27.
- Alifano P, Rivellini F, Piscitelli C, Arraiano CM, Bruni CB, Carlonagno MS. 1994. *Genes Dev.* 8:3021–31.
- Kurz JC, Nirajananakumari S, Fierke CA. 1998. *Biochemistry* 37:2393–400.
- Nirajananakumari S, Stains T, Cray SM, Christianson DW, Fierke CA. 1998. *Proc. Natl. Acad. Sci. USA* 95:15212–17.
- Gopalan V, Baxevanis AD, Landsman D, Altman S. 1997. *J. Mol. Biol.* 267:818–29.
- Hartmann RK, Heinrich J, Schlegl J, Schuster H. 1995. *Proc. Natl. Acad. Sci. USA* 92:5822–26.
- Chamberlain JR, Pagan R, Kindelberger DW, Engelke DR. 1996. *Nucleic Acids Res.* 24:3158–66.
- Schon A. 1999. *FEBS Microbiol. Rev.* 23:39–406.
- Oh BK, Pace NR. 1994. *Nucleic Acids Res.* 22:4087–94.
- Stains T, Nirajananakumari S, Fierke CA, Christianson DW. 1998. *Science* 280:752–55.
- Spitzfaden C, Nicholson N, Jones JJ, Guth S, Lehr R, et al. 2000. *J. Mol. Biol.* 295:105–15.
- Crary SM, Nirajananakumari S, Fierke CA. 1998. *Biochemistry* 37:9409–16.
- Ziehler WA, Day JJ, Fierke CA, Engelke DR. 2000. *Biochemistry* 39:9909–16.
- Lee Y, Kindelberger DW, Lee JY, McClellan S, Chamberlain J, Engelke DR. 1997. *RNA* 3:175–85.
- Maria RA, Intine RV. 2001. *Mol. Cell. Biol.* 21:367–79.
- Wolin SL, Matera AG. 1999. *Genes Dev.* 13:1–10.
- Pannone BK, Xue D, Wolin SL. 1998. *EMBO J.* 17:7442–53.
- Yoo CJ, Wolin SL. 1997. *Cell* 89:393–402.
- Mildvan AS. 1997. *Proteins* 29:401–16.
- Siecz TA, Sieitz JA. 1993. *Proc. Natl. Acad. Sci. USA* 90:6498–502.
- Beebe JA, Fierke CA. 1994. *Biochemistry* 33:10294–304.
- Pfeiffer T, Tekos A, Warnecke JM, Drainas D, Engelke DR, et al. 2000. *J. Mol. Biol.* 298:559–65.
- Thomas BC, Li XQ, Gegenheimer P. 2000. *RNA* 6:545–53.
- Thomas BC, Chamberlain J, Engelke DR. Gegenheimer P. 2000. *RNA* 6:554–62.

44. Tekos A, Tsagia A, Stathopoulos C, Drains D. 2000. *FEBS Lett.* 485:71-75.

45. Kurz JC, Fierke CA. 2000. *Curr Opin. Chem. Biol.* 4:553-58.

46. Beche JA, Kurz JC, Fierke CA. 1996. *Biochemistry* 35:10493-505.

47. Lee J-Y. 1991. *Characterization of Saccharomyces cerevisiae nuclear RNase P and its RNA subunit*. PhD thesis. Univ. Mich., Ann Arbor. 133 pp.

48. Altman S, Gopalan V, Vioque A. 2000. *RNA* 6:1689-94.

49. Gegenheimer P. 2000. *RNA* 6:1695-97.

50. Rossmanith W, Potuschak T. 2001. *Mol. Cell. Biol.* 21:8236-37.

51. Salavati R, Panigrahi AK, Stuart KD. 2001. *Mol. Biochem. Parasitol.* 115:109-17.

52. Miller DL, Martin NC. 1983. *J. Cell. Sci.* 34:91-17.

53. Underbrink-Lyon K, Miller DL, Ross NA, Fukuhara H, Martin NC. 1983. *Mol. Gen. Genet.* 191:512-18.

54. Morales MI, Dang YL, Lou YC, Sulo P, Martin NC. 1992. *Proc. Natl. Acad. Sci. USA* 89:9875-79.

55. Dang YL, Martin NC. 1993. *J. Biol. Chem.* 268:19791-96.

56. Shu HH, Wise CA, Clark-Walker GD, Martin NC. 1991. *Mol. Cell. Biol.* 11:1662-67.

57. Wise CA, Martin NC. 1991. *J. Biol. Chem.* 266:19154-57.

58. Wise CA. 1991. *Structural and functional studies of the yeast mitochondrial RNase P RNA*. PhD thesis. Univ. Tex. Southwestern, Dallas. 185 pp.

59. Stribinskis V, Gao GJ, Sulo P, Dang YL, Martin NC. 1996. *Mol. Cell. Biol.* 16:3429-36.

60. Stribinskis V, Gao GJ, Sulo P, Ellis SR, Martin NC. 2001. *Nucleic Acids Res.* 29:3631-37.

61. Kassenbrock CK, Gao GJ, Groom KR, Sulo P, Douglas MG, Martin NC. 1995. *Mol. Cell. Biol.* 15:4763-70.

62. Rossmanith W, Tullo A, Potuschak T, Karwan R, Sbisa E. 1995. *J. Biol. Chem.* 270:12885-91.

63. Rossmanith W, Karwan RM. 1998. *FEBS Lett.* 433:269-74.

64. Rossmanith W, Karwan RM. 1998. *Biochem. Biophys. Res. Commun.* 247:234-41.

65. Puranam RS, Attardi G. 2001. *Mol. Cell. Biol.* 21:548-61.

66. Gegenheimer P. 1995. *Mol. Biol. Rep.* 22:147-50.

67. Chen Y, Li XQ, Gegenheimer P. 1997. *Biochemistry* 36:2425-38.

68. Cordier A, Schon A. 1999. *J. Mol. Biol.* 289:9-20.

69. Baum M, Cordier A, Schon A. 1996. *J. Mol. Biol.* 257:43-52.

70. Pascual A, Vioque A. 1999. *FEBS Lett.* 442:7-10.

71. Collins IJ, Moulton V, Penny D. 2000. *J. Mol. Evol.* 5:194-204.

72. True HL, Celander DW. 1996. *J. Biol. Chem.* 271:16559-66.

73. Bantkiewicz M, Gold H, Altman S. 1989. *Genes Dev.* 3:488-99.

74. Doria M, Carrara G, Calandra P, Tocchini-Valentini GP. 1991. *Nucleic Acids Res.* 19:2315-20.

75. Kline L, Nishikawa S, Soll D. 1981. *J. Biol. Chem.* 256:5058-63.

76. Stathopoulos C, Kalpaxis DL, Draines D. 1995. *Eur. J. Biochem.* 228:976-80.

77. Jayanthi GP, Van Tuyl GC. 1992. *Arch. Biochem. Biophys.* 296:264-70.

78. Lee JY, Engelke DR. 1989. *Mol. Cell. Biol.* 9:2536-43.

79. Lee JY, Rohman CE, Molony LA, Engelke DR. 1991. *Mol. Cell. Biol.* 11:721-30.

80. Lygerou Z, Mitchell P, Petfalski E, Seraphin B, Tollervey D. 1994. *Genes Dev.* 8:1423-33.

81. Dichtl B, Tollervey D. 1997. *EMBO J.* 16:417-29.

82. Chu S, Zengel JM, Lindahl L. 1997. *RNA* 3:382-91.

83. Stole V, Altman S. 1997. *Genes Dev.* 11:2926-37.

84. Stole V, Katz A, Altman S. 1998. *Proc. Natl. Acad. Sci. USA* 95:6716-21.

85. Eder PS, Guerrier-Takada R, Stole V, Altman S. 1997. *Proc. Natl. Acad. Sci. USA* 94:1101-6.

86. Lygerou Z, Pluk H, van Venrooij WJ, Seraphin B. 1996. *EMBO J.* 15:5936-48.

87. van Eenennaam H, Pruijn GJM, van Venrooij WJ. 1999. *Nucleic Acids Res.* 27:2465-72.

88. Jarrous N, Eder PS, Wesolowski D, Altman S. 1999. *RNA* 5:153-57.

89. van Eenennaam H, Lugtenberg D, Vogelzangs JHP, van Venrooij WJ, Pruijn GJM. 2001. *J. Biol. Chem.* 276:31635-41.

90. Jarrous N, Reiner R, Wesolowski D, Mann H, Guerrier-Takada C, Altman S. 2001. *RNA* 7:1153-64.

91. Lee JY, Evans CF, Engelke DR. 1991. *Proc. Natl. Acad. Sci. USA* 88:6986-90.

92. Salgado-Garido J, Bragado-Nilsson E, Kandels-Lewis S, Seraphin B. 1999. *EMBO J.* 18:3451-62.

93. Tranguch AJ, Engelke DR. 1993. *J. Biol. Chem.* 268:14045-55.

94. Zimmerly S, Gamulin V, Burkard U, Soll D. 1990. *FEBS Lett.* 271:189-93.

95. Pitulle C, Garcia-Patis M, Zamudio KR, Pace NR. 1998. *Nucleic Acids Res.* 26:3333-39.

96. Altman S, Wesolowski D, Puranam RS. 1993. *Genomics* 18:418-22.

97. Martin NC, Lang BF. 1997. *Nucleic Acids Symp. Ser.* 36:42-44.

98. Kazakov S, Altman S. 1991. *Proc. Natl. Acad. Sci. USA* 88:9193-97.

99. Dar SC, Zito K, Smith D, Pace NR. 1992. *Biochemistry* 31:328-33.

100. Guerrier-Takada C, Altman S. 1993. *Biochemistry* 32:7152-61.

101. Haas ES, Brown JW, Piuille C, Pace NR. 1994. *Proc. Natl. Acad. Sci. USA* 91:2527-31.

102. Harris ME, Pace NR. 1995. *RNA* 1:210-18.

103. Pagan-Ramos E, Lee Y, Engelke DR. 1996. *RNA* 2:441-51.

104. Pagan-Ramos E, Lee Y, Engelke DR. 1996. *RNA* 2:1100-9.

105. Ziebler WA, Yang J, Kurochkin AV, Sandusky PO, Zuiderveld ER, Engelke DR. 1998. *Biochemistry* 37:3549-57.

106. Pagan-Ramos E, Tranguch AJ, Kindelberger DW, Engelke DR. 1994. *Nucleic Acids Res.* 22:200-7.

107. Loria A, Pan T. 1996. *RNA* 2:551-63.

108. Loria A, Pan T. 2001. *Nucleic Acids Res.* 29:1892-97.

109. Massire C, Jager L, Westhof E. 1998. *J. Mol. Biol.* 279:773-93.

110. Fang XW, Pan T, Sosnick TR. 1999. *Biochemistry* 38:16840-46.

111. Mobley EM, Pan T. 1999. *Nucleic Acids Res.* 27:4298-304.

112. Ciesielska J, Hardt WD, Schlegl J, Edmann VA, Hartmann RK. 1994. *Eur. J. Biochem.* 219:49-56.

113. Burgin AB, Pace NR. 1990. *EMBO J.* 9:4111-18.

114. Harris ME, Nolan JM, Malhotra A, Brown JW, Harvey SC, Pace NR. 1994. *EMBO J.* 13:3953-63.

115. Morse DP, Schmidt EJ. 1993. *J. Mol. Biol.* 230:11-14.

116. Ziebler WA, Morris J, Scott FH, Millikin C, Engelke DR. 2001. *RNA* 7:565-75.

117. Yuan Y, Tan E, Reddy R. 1991. *Mol. Cell. Biol.* 11:5266-74.

118. Liu MH, Yuan Y, Reddy R. 1994. *Mol. Cell. Biochem.* 130:75-82.

119. Jacobson MR, Cao LG, Taneja K, Singer RH, Wang YL, Pederson T. 1997. *J. Cell. Sci.* 110:829-37.

120. Kirsebom LA, Svärd SG. 1994. *EMBO J.* 13:4870-76.

121. Glemarec C, Kufel J, Voldesi A, Maltseva T, Sandstrom A, et al. 1996. *Nucleic Acids Res.* 24:2022-35.

122. Kufel J, Kirschbom LA. 1998. *RNA* 4:777-88.

123. Oh BK, Frank DN, Pace NR. 1998. *Biochemistry* 37:7277-83

124. Tuerk C, Gauss P, Thernes C, Grobe DK, Gayle M, et al. 1988. *Proc. Natl. Acad. Sci. USA* 85:1364-68

125. Cheong C, Varani G, Timcoo Jr. 1990. *Nature* 346:680-82

126. Heus HA, Pardi A. 1991. *Science* 253:191-94

127. Jaeger L, Michel F, Westhof E. 1994. *J. Mol. Biol.* 236:1271-76

128. Murphy FL, Cech TR. 1994. *J. Mol. Biol.* 236:49-63

129. Cate JH, Gooding AR, Podell E, Zhou K, Golden BL, et al. 1996. *Science* 273:1696-99

130. Gluck A, Endo Y, Wool IG. 1992. *J. Mol. Biol.* 226:411-24

131. Krummel DA, Altman S. 1999. *Proc. Natl. Acad. Sci. USA* 96:11200-5

132. Brusca EM, True HL, Celdander DW. 2001. *J. Biol. Chem.* 276:42543-48

133. Li Y, Altman S. 2001. *Proc. Natl. Acad. Sci. USA* 98:441-44

134. Jiang T, Altman S. 2001. *Proc. Natl. Acad. Sci. USA* 98:920-25

135. True HL, Celdander DW. 1998. *J. Biol. Chem.* 273:7193-96

136. Calvo O, Cuesta R, Anderson J, Gutierrez N, Garcia-Barrio MT, et al. 1999. *Mol. Cell. Biol.* 19:4167-81

137. Clark MW, Abelson J. 1987. *J. Cell. Biol.* 105:155-26

138. Sarkar S, Hopper AK. 1998. *Mol. Biol. Cell* 9:3041-55

139. Bertrand E, Houser-Scott F, Kendall A, Singer RH, Engelke DR. 1998. *Genes Dev.* 12:2463-68

140. Fields S, Song O. 1989. *Nature* 340:245-46

141. SenGupta DJ, Zhang B, Kraemer B, Pochart P, Fields S, Wickens M. 1996. *Proc. Natl. Acad. Sci. USA* 93:8496-501

142. Jiang T, Guenier-Takada C, Altman S. 2001. *RNA* 7:97-41

143. Chang DD, Clayton DA. 1987. *EMBO J.* 6:409-17

144. Chang DD, Clayton DA. 1987. *Science* 235:1178-84

145. Reimer G, Raska I, Scheer U, Tan EM. 1988. *Exp. Cell Res.* 176:117-28

146. Kiss T, Marshallsay C, Filipowicz W. 1992. *EMBO J.* 11:3737-46

147. Li K, Smagula CS, Parsons WI, Richardson JA, Gonzalez M, et al. 1994. *J. Cell Biol.* 124:871-82

148. Jacobson MR, Cao JG, Wang YL, Peterson T. 1995. *J. Cell Biol.* 131:1649-58

149. Tollervey D. 1995. *Mol. Biol. Rep.* 22:75-79

150. Lindahl L, Zengel JM. 1995. *Mol. Biol. Rep.* 22:69-73

151. Schmitt ME, Clayton DA. 1993. *Mol. Cell. Biol.* 13:7935-41

152. Lygerou Z, Allmang C, Tollervey D, Seraphin B. 1996. *Science* 272:268-70

153. Chu S, Archer RH, Zengel JM, Lindahl L. 1994. *Proc. Natl. Acad. Sci. USA* 91:659-63

154. Cai T, Reilly TR, Cerio M, Schmitt ME. 1999. *Mol. Cell. Biol.* 19:7857-69

155. Shiba E, Pesole G, Tullo A, Saccone C. 1996. *J. Mol. Evol.* 43:46-57

156. Schmitt ME, Clayton DA. 1994. *Genes Dev.* 8:2617-28

157. Forster AC, Altman S. 1990. *Cell* 62:407-9

158. Karwan R. 1993. *FEBS Lett.* 319:1-4

159. Schmitt ME, Bennett JL, Dairaghi DJ, Clayton DA. 1993. *FASEB J.* 7:208-13

160. Schmitt ME. 1999. *J. Mol. Biol.* 292:827-36

161. Lindahl L, Fretz S, Epps N, Zengel JM. 2000. *RNA* 6:653-58

162. Pluk H, van Fennema H, Ruijtes SA, Zengel JM, Lindahl L, Engelke DR. 2002. *Proc. Natl. Acad. Sci. USA* 99:2684-89

163. Kendall A, Hull MW, Bertrand E, Good PR, Singer RH, Engelke DR. 2000.